ED KIANKA
CONGRESSIONAL INFORMATION SVC
4520 EAST-WEST HIGHWAY
BETHESDA MD 20914
86389
028

American Journal of Public Health

December 1991, Vol. 81, No. 12 Established 1911

Editorials

1557 Strategies to Prevent HIV Infection in the United States *A. R. Hinman*

1990 Keynote Address

1560 Health for All: A Public Health Vision *W. H. McBeath*

Public Health Policy Forum

1566 Lessons from London: The British Are Reforming Their National Health Service *A. Vall-Spinosa*

Featuring High-Risk Behavior

1571 Injury-Related Medical Care Utilization in a Problem Drinking Population *J. O. Blose and H. D. Holder*

1576 HIV Infection, Genital Ulcer Disease, and Crack Cocaine Use among Patients Attending a Clinic for Sexually Transmitted Diseases *K. Chirgwin, J. A. DeHovitz, S. Dillon, and W. M. McCormack*

1580 Reduction of High-Risk Sexual Behavior among Heterosexuals Undergoing HIV Antibody Testing: A Randomized Clinical Trial *N. S. Wenger, L. S. Linn, M. Epstein, and M. F. Shapiro*

1586 Behavior Changes after Notification of HIV Infection *P. D. Cleary, N. Van Devanter, T. F. Rogers, E. Singer, R. Shipton-Levy, M. Steilen, A. Stuart, J. Avorn, and J. Pindyck*

1591 Knowledge about AIDS and HIV in the US Adult Population: Influence of the Local Incidence of AIDS *L. F. McCaig, A. M. Hardy, and D. M. Winn*

1596 HIV Instruction, HIV Knowledge, and Drug Injection among High School Students in the United States *D. Holtzman, J. E. Anderson, L. Kann, S. L. Arday, B. I. Truman, and L. J. Kolbe*

1602 Needle Sharing in The Netherlands: An Ethnographic Analysis *J. C. Grund, C. D. Kaplan, and N. F. P. Adriaans*

1636 Overcoming Denial and Increasing the Intention to Use Condoms through the Induction of Hypocrisy *E. Aronson, C. Fried, and J. Stone*

1638 Knowledge about HIV and Behavioral Risks of Foreign-Born Boston Public School Students *R. W. Hingson, L. Strunin, M. Grady, N. Strunk, R. Carr, B. Berlin, and D. E. Craven*

1642 HIV Infection and Risk Behaviors among Intravenous Drug Users in Low Seroprevalence Areas in the Midwest *H. A. Siegal, R. G. Carlson, R. Falck, L. Li, M. A. Forney, R. C. Rapp, K. Baumgartner, W. Myers, and M. Nelson*

1645 Frequency and Thoroughness of STD/HIV Risk Assessment by Physicians in a High-Risk Metropolitan Area *B. O. Boekeloo, E. S. Marx, A. H. Kral, S. C. Coughlin, M. Bowman, and D. L. Rabin*

(continued page 1543)

APHA
American Public Health Association

Featuring High-Risk Behavior Publishers' Issue

Contents

American Journal of Public Health

"Information for Authors" appears on page 134 in the January 1991 issue of the Journal and may also be obtained from the Journal Editorial Office. An abbreviated version is printed in this issue.

American Journal of Public Health
1015 Fifteenth Street, NW
Washington, DC 20005

American Public Health Association

December 1991, Vol. 81, No. 12 ISSN: 0090-0036 Established 1911

1648 Changes in AIDS Case Reporting after Hospital Site Visits *D. Fife, J. McAnaney, and M. A. Rahman*

Articles

1608 Socioeconomic Status and Electrolyte Intake in Black Adults: The Pitt County Study *A. M. Gerber, S. A. James, A. S. Ammerman, N. L. Keenan, J. M. Garrett, D. S. Strogatz, and P. S. Haines*

1613 Representativeness of Deaths Identified through the Injury-at-Work Item on the Death Certificate: Implications for Surveillance *J. Russell and C. Conroy*

1619 Occupational Health and Safety in Brazil *H. Frumkin and V. de M. Câmara*

1625 An Evaluation of the Impact of Maternity Care Coordination on Medicaid Birth Outcomes in North Carolina *P. A. Buescher, M. S. Roth, D. Williams, and C. M. Goforth*

1630 The Importance of Assessing the Fit of Logistic Regression Models: A Case Study *D. W. Hosmer, S. Taber, and S. Lemeshow*

Public Health Briefs

1651 Patterns of Breast and Cervical Cancer Screening at Three Public Health Centers in an Inner-City Urban Area *S. Whitman, D. Ansell, L. Lacey, E. H. Chen, N. Ebie, J. Dell, and C. W. Phillips*

1654 The Impact of Vitamin A Supplementation on Morbidity: A Randomized Community Intervention Trial *M. H. Abdeljaber, A. S. Monto, R. L. Tilden, M. A. Schork, and I. Tarwotjo*

1656 Occupational Injuries and Illnesses among Washington State Agricultural Workers *P. Demers and L. Rosenstock*

1659 The Effects of Removing a Statutory Barrier to Workers' Compensation for Farm Workers *P. Demers and L. Rosenstock*

Public Health and the Law

1661 Epidemiology and the Law: Courts and Confidence Intervals *T. Christoffel and S. P. Teret*

Other Journal Departments

1544 Masthead

1546 Advertisers' Index

1570 Erratum

1667 Notes from the Field

1670 New from NCHS at CDC

1673 Letters to the Editor

1681 Book Corner

1696 Information for Authors

1698 Job Opportunities

1706 Index

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.

APHA OFFICERS
C. William Keck, MD, MPH, *President*
Myron Allukian, Jr., DDS, MPH, *Immediate Past President*
Joyce C. Lashof, MD, *President-Elect*
James M. Dunning, DDS, MPH, *Vice-President (USA)*
Dr. Robert G. Evans, *Vice-President (Canada)*
Dr. Hernán Delgato, *Vice-President (Latin America)*
Caswell A. Evans, Jr., MD, MPH *Treasurer*
Shirley Randolph, MSPH, *Chair, Executive Board*
George E. Hardy, Jr., MD, MPH, *Speaker, Governing Council*
William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD
Shirley Randolph, MSPH (1991), *Chair*
Marilyn Aguirre-Molina, MS, EdD (1993)
Myron Allukian, Jr., DDS, MPH, *Immediate Past President*
Molly Joel Coye, MD, MPH (1991)
Brian A. Dobrow, MPH (1992)
Caswell A. Evans, Jr., DDS, MPH, *Treasurer*
Eugene Feingold, PhD (1991)
Kristine M. Gebbie, BSN, MN (1994)
Janet Heinrich, MPH, DrPH (1993)
Margaret H. Jordan, RN, MPH (1992)
C. William Keck, MD, MPH, *President*
Joyce C. Lashof, MD, *President-Elect*
Barry S. Levy, MD, MPH (1993)
Allan Rosenfield, MD (1994)
Reed V. Tuckson, MD (1994)
Frances M. Veverka, MPH, RS (1992)

EX OFFICIO
William H. McBeath, MD, MPH, *Executive Director*
Patricia P. Evans, MPH, *Chair, Action Board*
George Pickett, MD, MPH, *Chair, Program Development Board*
George E. Hardy, Jr., MD, MPH, *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc., 1015 Fifteenth St., NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50,000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Michel A. Ibrahim, MD, PhD, Editor, *American Journal of Public Health,* 1015 Fifteenth St., NW, Washington, DC 20005.

Information on Manuscript Preparation is published in this Journal in January. The Journal publishes reports of original research, demonstrations, evaluations, and other articles covering current aspects of public health. Manuscripts are accepted for consideration with the understanding that they, or their essential substance, have not been published previously or accepted for publication and are not under consideration elsewhere.

Subscription price, individual, is $80/year US, $120/year international; **institutional,** $160/year US, $240/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to American Journal of Public Health, 1015 Fifteenth St., NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact: Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Kammann, Advertising Manager, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal of APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact: Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols. 1 to present are also available on microfilm.

Copyright © 1991 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS Bibliographic Retrieval Service Inc, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc., Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Injury-Related Medical Care Utilization in a Problem Drinking Population

James O. Blose, MPP, and Harold D. Holder, PhD

ABSTRACT

Background. Population-based data on the overall risk of injury among problem drinkers are extremely limited.

Methods. We conducted an eight-year study of injury-related medical care utilization for a group of problem drinkers (n = 3 729) enrolled in the health insurance plan of a large manufacturing firm with plants in a number of midwestern states. A comparison non-problem drinking cohort matched on age, gender, and number of years of enrollment with the same company plan was also utilized.

Results. Problem drinkers utilized injury-related medical care at a rate 1.6 times that of the comparison group and experienced injury-related medical care costs which were three times as high. Increased risks were found for both men and women and for all ages studied.

Conclusions. This study is one of the few to examine the overall incidence of fatal and non-fatal injuries among a large population of problem drinkers and provides evidence that higher medical care costs are associated with chronic alcohol use. (*Am J Public Health* 1991;81:1571–1575)

Introduction

The role of injuries as a public health problem and as a contributor to medical care costs among the general population has seen increasing attention.[1,2] The association of alcohol use with accident and injury has also been the subject of frequent investigation. Yet there is surprisingly little epidemiological data regarding the overall risk of injury—and the attendant utilization of medical care—among alcoholics and problem drinkers.

A considerable body of research has examined the relationship between alcohol use and fatal injuries. (See reviews by Smith and Kraus,[3] Howland and Hingson,[4,5] Hingson and Howland.[6]) The association between alcohol and traffic fatalities has received particular scrutiny.[7,8] On the whole, this research suggests that there is indeed a link between alcohol and fatal injury, although the actual magnitude of the risk cannot be estimated with great precision.

Fatal injuries, however, constitute only a small portion of all injuries.[2,9] The connection between alcohol and nonfatal injuries has received much less attention. Most research on nonfatal injuries focuses solely on traffic accidents.[10,11] Other studies of nonfatal injuries (i.e. those not limited to traffic accidents) are entirely event based—that is, they are based on the analysis of data from accident victims presenting themselves for care at specific facilities such as emergency rooms.[12–17]

While these event-based studies provide important data on the association between drinking and injury, they provide us with little information on the overall incidence of nonfatal injuries among problem drinkers as a group. Roizen[18] reviewed the literature through 1979 and found no studies on nonfatal accidents (other than traffic accidents) which were based on populations of problem drinkers or alcoholics. Our review of the more recent literature has likewise revealed no such studies.

Several national studies of self-reported alcohol-related problems have been published (e.g. Hilton and Clark[19]). The single survey-based study reporting specific data on nonfatal injuries (Hingson, *et al.*[20]) found that high levels of typical alcohol consumption were associated with a higher level of accidental injury in employee populations in four New England states. Such surveys, however, cannot identify diagnosed problem drinkers and alcoholics.

Methods

This study derives estimates of the incidence of medical care utilization due to injury for a problem drinking population. Health insurance claims and enrollment data from a large midwestern manufacturing corporation were analyzed for the years 1974 through 1987. The data base included approximately 20 million health care claims filed by the 260,000 enrollees covered during the 14-year period.

For the 1980–87 period over which ICD-9 (International Classification of Diseases, 9th Revision) codes were available, 3,729 individuals received a diagnosis in-

Address reprint requests to James O. Blose, MPP, Pacific Institute for Research and Evaluation, 121 West Rosemary Street, 2nd floor, Chapel Hill, NC 27516. Dr. Holder is with the Prevention Research Center, Pacific Institute for Research and Evaluation, Berkeley, CA. This paper, submitted to the Journal May 24, 1990, was revised and accepted for publication January 23, 1991.

TABLE 1—Injury-related Medical Care: Average Annual Number of Events, Cost and Inpatient Days for Problem-drinking Group and Matched Group

Number of Injury-Related Medical Care Events	Problem Drinkers N = 3729	Matches N = 3729	P Value[a]	Risk Estimate[b]
All events	1.32 (±.08)	.76 (±.06)	.002	1.74
Inpatient admissions	.04 (±.004)	.01 (±.002)	.001	4.00
Outpatient visits	1.05 (±.07)	.65 (±.06)	.002	1.62
Emergency room admissions	.23 (±.01)	.10 (±.01)	.001	2.30
Total Cost of All Injury-Related Medical Care Events[c]	$250 (±26)	$82 (±11)	<.001	3.05
Average Number of Injury-Related Inpatient Days	.32 (±.03)	.08 (±.04)	<.001	4.00

Annual averages per person, 95% confidence intervals in parentheses.
[a]Based on T test.
[b]Ratio of observed to expected.
[c]December 1985 dollars.

dicating a chronic drinking problem.* A group of non-problem drinking individuals, i.e. individuals who had never received a diagnosis indicating a chronic drinking problem, was selected from the total enrollee population. These individuals were matched on a one-to-one basis with members of the problem drinker group on age, gender, and the number of years enrolled. The average age of both groups was 38.6 years of age; 75.5 percent of each group was male.

Coverage under the company's broad health insurance policy was quite stable due to the largely unionized nature of the workforce and to the fact that the company's program was self-insured and administered throughout the 14-year period. Company-sponsored HMO options were established in the early 1980s, with benefit provisions similar to the fee-for-service plans. Utilization records for the two major HMOs were available while records for a third HMO were not. Further details on the company's insurance coverage and data base are described elsewhere.**

The diagnosis codes from the Clinical Modification of the 9th Revision of the International Classification of Diseases[21] were used to classify events as injury-related or non injury-related. Events carrying a diagnosis of 800–995 (the injury and poisoning grouping) were classified as injury-related. The supplementary codes for classification of external causes of injury and poisoning were not available. Among the problem drinking group, 9.1 percent also exhibited evidence of the abuse of other drugs—ICD-9 diagnoses of drug psychoses (292), drug dependence (304), or drug abuse (305.2–305.9).

For the two study populations (total N = 7,458) over 50,000 injury-related medical care events (admissions and/or visits) were identified during the 1980–87 period. Measures were developed for inpatient, outpatient, and emergency room events. Each visit or admission was considered a separate event. The actual number of unique injuries cannot be ascertained from the data, nor can events related to fatal and nonfatal injuries be readily distinguished. No data were available on whether drinking immediately preceded a given event.

Data were also available on the total cost (defined as total charges) for injury-related medical care events. The Medical Care Index (US Department of Labor) was used to adjust all dollar values (reported in December 1985 dollars) for inflation.

All annual averages were calculated based upon the number of months of enrollment (months of exposure) for each individual. Individuals who switched into the HMO plan for which data were not available were considered to have lost their enrollment eligibility at that point. On average, 49 percent of all injury-related medical care events occurred prior to the initial problem-drinking diagnosis, while 51 percent occurred afterward. The average length of enrollment was about six years, approximately equally divided into pre- and post-diagnosis periods.

One class of injuries is underrepresented in the data base. On-site medical staff at major company worksites generally provide medical care for minor on-the-job injuries. No insurance claims are filed if all care is provided on site by company medical personnel. Medical care for work injury was included in our data only if off-site medical services were used.

Selection bias could occur if the manner in which study group members are chosen has a differential impact on the dependent measures for the two groups. Selection into the problem-drinking group is dependent on receiving a specific chronic disease diagnosis while membership in the control group is not. Since the dependent measures used for both groups included injury-related medical care only, selection bias is most likely if a chronic drinking diagnosis is directly associated with an injury-related event. However, only 7 percent of all problem drinkers received any injury-related medical care in the month preceding their initial problem-drinking diagnosis. Thus, there is no evidence that a significant selection bias exists for the study groups used in the analyses reported here.

Results

Table 1 presents the average incidence rates for injury-related medical care events over the 1980–87 period. Problem drinkers averaged 1.32 medical care events per year compared to only .76 per year for those with no chronic drinking diagnosis. The difference between the two groups is especially pronounced for events involving hospitalization. For both groups, injury-related outpatient care events far outnumber emergency room and inpatient admissions. The problem-drinking group also experienced significantly higher injury-related medical costs than did their matched cohort—an average of $250 per year compared to $82. The number of injury-related inpatient days (number of days of hospitalization due to injury) was .32 and .08 per year for the two groups, respectively.

The incidence of injury-related medical care utilization among problem drinkers exceeds the incidence for matches for all

*The following ICD-9 codes were used[21]:
291 —alcoholic psychoses
303 —alcohol dependence syndrome
305.0—alcohol abuse
357.5—alcoholic polyneuropathy
425.5—alcoholic cardiomyopathy
535.3—alcoholic gastritis
571 —chronic liver disease and cirrhosis
572.3—portal hypertension

**Holder HD, Blose JO: The reduction of health care costs associated with alcoholism treatment: A 14-year longitudinal study. J Stud Alcohol (in press).

demographic subgroups examined (Table 2). All differences were statistically significant. Males have much higher injury care rates than females in both groups. The highest injury rates occurred among those ages 31–50, while those age 51 and over showed comparatively low rates of injury.

Both male and female problem drinkers incur annual injury-related medical care costs which are three times the cost of their matches (Table 3). For all subgroups, the risk estimates for costs (Table 3) exceed those for total injury events (Table 2). This may be due to a higher incidence of severe injuries among problem drinkers or to variations in the type of injury. No data are available on these factors.

Adjusting for Demographic Differences between Groups

While the problem drinker and non-problem drinker groups are matched on age, gender, and length of enrollment, other demographic differences between the groups exist (see Table 4). Analysis of variance was used to control for the impact of these demographic factors on the total number of injury-related medical care events. Demographic control variables include GENDER, AGE (30 and under, 31–50, and 51 and over); work class (WORKCL), whether an individual is a salaried or hourly employee; PERSTYPE (i.e. whether an individual is an employee, spouse, child); insurance plan type (PLANTYPE). GROUP denotes whether an individual is in the problem drinking or matched group. All independent variables except gender were statistically significant and were thus included in the final model (Table 5). No interaction terms were statistically significant. The model indicates that after the impact of demographic variables is considered, the difference in injury-related medical care utilization between problem drinkers and matches is still statistically significant ($p = .002$).

Overall, the adjusted level of injury-related medical care events for problem drinkers is 1.6 times what would otherwise be expected (Table 6). While female problem drinkers experience a lower overall level of injury-related care utilization than males, their relative risk is somewhat higher (1.91 vs 1.54). Similarly, problem drinkers over age 51 experience the lowest incidence but the highest relative risk (2.12).

TABLE 2—Average Annual Number of Injury-related Medical Care Events by Gender and Age

	Problem Drinkers	Matches	P Value[a]	Risk Estimate[b]
Gender				
Male	1.46 (±.09)	.82 (±.07)	<.001	1.78
Female	.88 (±.13)	.58 (±.14)	.002	1.52
Age (years)				
30 and under	1.34 (±.12)	.65 (±.08)	<.001	2.06
31–50	1.62 (±.15)	.96 (±.11)	<.001	1.69
51+	.75 (±.10)	.57 (±.08)	.023	1.32

95% confidence intervals in parentheses.
[a]T test.
[b]Ratio of observed to expected.

TABLE 3—Average Annual Cost[a] of Injury-related Medical Care by Gender and Age

	Problem Drinkers	Matches	P Value[b]	Risk Estimate[c]
Gender				
Male	$263 (±30)	$86 (±13)		3.06
Female	$210 (±50)	$67 (±17)		3.13
Age (years)				
30 and under	$255 (±40)	$67 (±13)	<.001	3.81
31–50	$248 (±39)	$74 (±10)	<.001	3.35
51+	$247 (±62)	$116 (±36)	<.001	2.13

95% confidence intervals in parentheses.
[a]December 1985 dollars.
[b]T test
[c]Ratio of observed to expected.

TABLE 4—Demographic Comparison: Problem Drinkers and Matches

	Problem Drinkers	Matches
Person Type (PERSTYPE)		
% employee	61.3	56.0
% spouse	18.7	23.2
% child	19.9	20.8
Work Class (WORKCL)[a]		
% hourly	80.8	62.2
% salaried	19.2	37.8
Insurance Plan Type (PLANTYPE)		
% traditional only	53.6	43.8
% HMO only	-0-	.2
% traditional/HMO[b]	46.4	56.0

[a]Individuals who are not company employees are classified according to the work class of the employed family member.
[b]Switched into HMO from traditional fee-for-service plan.

Discussion

This is a unique population-based study comparing overall injury rates among problem drinkers and a matched age and gender cohort. During the eight-year period from 1980–87, the average number of injury-related medical care events per year for over 3,700 problem drinkers exceeded that for the matched group for both males and females and for all age groups. Relative risk estimates range from 1.54 for males to 2.12 for those age 51 and over, based on an analysis of variance model which adjusts for differences due to person type, work class, and insurance plan type. These findings provide clear evidence that the number of medical care events due to injury is significantly higher for problem drinkers than for demographically similar individuals who do not have a known history of problem drinking.

TABLE 5—Final Analysis of Variance Model

	Sum of Squares	DF	Mean Square	F Ratio	P
GROUP	318.84	1	318.84	65.48	.002
AGE	519.33	2	259.66	53.32	<.001
PERSTYPE	739.39	2	369.70	75.92	.003
WORKCL	164.03	1	164.03	33.69	.002
PLANTYPE	30.24	2	15.12	3.11	.025
Error	34919.54	7171	4.87		

TABLE 6—Number of Injury-related Medical Care Events for Problem Drinkers—Adjusted Annual Means[a] and Risk Estimates[b]

	Problem Drinker Group	Matched Group	Risk Estimate
All individuals	1.32	.82	1.61
Males	1.42	.92	1.54
Females	1.01	.53	1.91
Age (years)			
30 and under	1.29	.78	1.65
31–50	1.58	1.07	1.48
51+	.91	.43	2.12

[a]Adjusted mean number of injury-related medical care events per year, based on ANOVA results.
[b]Ratio of observed to expected.

This study could not provide direct information on the actual risk of injuries among problem drinkers since the data contained information only on medical care, and not on specific accidents or injuries. In fact, our estimates of the number of injury-related medical events (1.32 per year for problem drinkers; .76 for matches) are significantly higher than rates of self-reported injury reported by survey-based studies of the general population. Hingson, *et al*,[20] in a survey of 1,740 employed adults in New England, found a self-reported annual rate of .22 injuries per person. Data from the National Health Survey[22] indicate a per person annual rate of injury of .38 for persons ages 17–44 and .22 for persons over age 44.

There are several possible reasons for this difference. First, some injuries undoubtedly result in more than one episode of medical care, particularly in cases of serious injuries where extensive follow-up care may be required. Secondly, the number of self-reported injuries may be an underestimate of the actual number. How accurately survey respondents report injuries (usually defined as injuries requiring medical attention) occurring over the course of a year is not clear.

In the end, the observed differences between our problem drinking and nonproblem drinking groups are consistent with other research. Hingson, *et al*,[20] found that persons whose self-reported intake of alcoholic beverages averaged five or more drinks per day had a relative risk of self-reported accidental injury of 1.7, compared with abstainers. This is quite similar to our estimate of the relative risk of injury-related medical care for problem drinkers (1.61).

Our results clearly imply that problem drinkers have either higher overall rates of injury, injuries which are more serious, or both. It is possible that the increased risk among problem drinkers is not due solely to their higher rate of alcoholic beverage consumption. Heavy drinkers may, for instance, be more likely to engage in activities which pose a greater risk of injury. It is also possible that the chronic use of alcohol impedes the process of recovery from physical injury. Additional research will be required to identify the role of these factors.

These findings provide important evidence that higher cost and utilization of medical care related to injury is associated with chronic alcohol use, supporting the view that prevention and treatment of alcohol abuse may be an important approach to reducing injury costs. □

Acknowledgments

Support for this paper was provided by the National Institute on Alcohol Abuse and Alcoholism, grant # RO1 AA06248, "Health Care Costs for Employed Alcoholics in Treatment."

References

1. Harlan LC, Harlan WR, Parsons PE: The economic impact of injuries: A major source of medical costs. Am J Public Health 1990; 80:453–459.
2. Rice DP, MacKenzie EJ, and associates (eds): Cost of Injury in the United States: A Report to Congress. San Francisco: Institute for Health & Aging, University of California, and Injury Prevention Center, Johns Hopkins University, 1989.
3. Smith GS, Kraus JF: Alcohol and residential, recreational, and occupational injuries: A review of the epidemiologic evidence. Annu Rev Public Health 1988; 9:99–121.
4. Howland J, Hingson R: Alcohol as a risk factor for injuries or death due to fires and burns: Review of the literature. Public Health Rep 1987; 102:475–483.
5. Howland J, Hingson R: Alcohol as a risk factor for drownings: A review of the literature (1950–1985). Accid Anal & Prev 1988; 20:19–25.
6. Hingson R, Howland J: Alcohol as a risk factor for injury or death resulting from accidental falls: A review of the literature. J Stud Alcohol 1987; 48:212–219.
7. US Department of Transportation, National Highway Traffic Safety Administration: Alcohol and Highway Safety 1984: A Review of the State of Knowledge. DOT Pub. No. DOT-HS-806-569. Washington, DC: Govt Printing Office, 1985.
8. Cerrelli EC: Alcohol in Fatal Accidents, National Estimates-USA. DOT Pub. No. DOT-HS-806 371. Springfield, VA: National Technical Information Service, 1983.
9. Committee on Trauma Research, Commission on Life Sciences, National Research Council, and the Institute of Medicine: Injury in America: A Continuing Public Health Problem. Washington, DC: National Academy Press, 1985.
10. Borkenstein RF, Crowther RF, Shumate RP, Ziel WB, Zylman R: The Role of the Drinking Driver in Traffic Accidents. Bloomington, IN: Indiana University, Dept of Police Administration, 1964.
11. US Department of Transportation, National Highway Traffic Safety Administration: National Accident Sampling System 1983: A Report on Traffic Accidents and Injuries in the US. DOT Pub. No. DOT-HS-806 699. Springfield VA: National Technical Information Service, 1985.
12. Wechsler H, Kasey EH, Thum D, Demone HW: Alcohol level and home accidents. Public Health Rep 1969; 84:1043–1050.
13. Honkanen R, Ertama L, Kuosmanen P, Linnoila M, Alha A, Visuri T: The role of alcohol in accidental falls. J Stud Alcohol 1983; 44:231–245.
14. Stephens CJ: Alcohol consumption and casualties: Drinking in the event. Drug Alcohol Dependence 1987; 20:115–127.
15. Cherpitel CJS: Drinking patterns and problems associated with injury status in emergency room admissions. Alcohol Clin Exp Res 1988; 12:105–110.

16. Cherpitel CJS: Prediction of alcohol-related casualties among emergency room admissions. Int J Addict 1989; 24:725–737.
17. Cherpitel CJS: Prediction of alcohol-related casualties: A comparison of two emergency room populations. Drug Alcohol Dependence 1989; 24:195–203.
18. Roizen J: Estimating alcohol involvement in serious events. *In:* National Institute on Alcohol Abuse and Alcoholism: Alcohol and Health Monograph 1, Alcohol Consumption and Related Problems. PB87-106803. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism, 1982; 179–222.
19. Hilton ME, Clark WB: Changes in American drinking patterns and problems, 1967–1984. J Stud Alcohol 1987; 48:515–522.
20. Hingson RW, Lederman RI, Walsh DC: Employee drinking patterns and accidental injury: A study of four New England states. J Stud Alcohol 1985; 46:298–303.
21. Commission on Professional and Hospital Activities: ICD-9-CM: International Classification of Diseases, 9th Rev. Vol. 1. Ann Arbor, MI: The Commission, 1980.
22. US Dept of Health and Human Services: Types of Injuries and Impairments due to Injuries, United States. Data from the National Health Survey, Series 10, No. 159. DHHS Pub. No. (PHS) 87-1587. Hyattsville, MD: National Center for Health Statistics, 1986.

Jenny Labalme, Anniston Star *Medical Writer, Wins 21st Annual Ray Bruner Science Writing Award*

Jenny Labalme, Medical Writer for the *Anniston Star,* is the 1991 recipient of the Ray Bruner Science Writing Award, which is given annually by the American Public Health Association. Labalme was chosen for her strong investigative skills, her sense for important issues, her thorough and thoughtful coverage, and her ability to combine compassion with objectivity.

In a series of several articles, Labalme investigated length-of-stay issues at the psychiatric unit of Anniston's Regional Medical Center (RMC). Former and current employees charged that the clinic kept patients the maximum number of days their insurance would cover regardless of medical need. Labalme's stories spurred still ongoing investigations by Medicare and Blue Cross/Blue Shield. "Jenny wrote these stories before ABC's Prime Time Live followed with a similar series on other psychiatric units run by the same organization elsewhere in the nation," said Joe Distelheim, Executive Editor of the *Anniston Star.*

Since Labalme reported on the alleged insurance milking at the local RMC, Psychiatric Institute, which operates the local clinic in conjunction with RMC, has come under scrutiny by the national media and law enforcement agencies in at least three other states, with particularly serious charges of fraud being investigated in Texas.

In another of her winning entries, a story on pesticides, Labalme provided detailed information on the hazards of the 10 most widely used household pesticides and the need for more aggressive involvement by the EPA. A story on Lay Lake carefully examined the issues surrounding a state health department's announcement to lift a ban on dioxin-tainted fish. According to Distelheim, Labalme's stories on the state's first rural AIDS clinic and the medical community's struggle to deal with their own ambivalent feelings toward AIDS patients has "stirred debate and soul-searching" among the 34,000 readers of the *Anniston Star.*

A panel of five distinguished judges selected the winner for this year's Bruner Award: Tim Friend, Health/Science Editor, *USA Today;* Maureen Mylander, Medical Editor/Writer for NIH's Office of Communications and author of *The Healthy Male—A Comprehensive Health Guide for Men;* Ronald J. Sanchez, Editor, *American Medical Writers Association Journal;* Lynn Wagner, Washington Bureau Chief, *Modern Healthcare;* and Emanuel Landau, PhD, Scientific Consultant, American Public Health Association.

The Ray Bruner Science Writing Award was established in 1971 to recognize outstanding achievement in news reporting by young journalists in the health/science/medical reporting field. It is administered by the American Public Health Association and sponsored by Lederle Laboratories, a division of American Cyanamid Company.

The award commemorates Ray Bruner, a former science editor of the *Toledo Blade* (Ohio), who dedicated much of his time during his 40-year career to encouraging young journalists to enter and excel in the science writing field.

HIV Infection, Genital Ulcer Disease, and Crack Cocaine Use among Patients Attending a Clinic for Sexually Transmitted Diseases

Keith Chirgwin, MD, Jack A. DeHovitz, MD, MPH, Stephen Dillon, MD, and William M. McCormack, MD

ABSTRACT

Background. Recently there has been a rise in genital ulcer disease (GUD) in urban minority heterosexuals in the United States. The impact of these increased GUD rates on HIV transmission patterns in this population is unknown.

Methods. Sexually transmitted disease (STD) diagnoses were correlated with HIV antibody status and risk factor history in 194 patients who consented to HIV testing at an STD clinic in central Brooklyn.

Results. Of 36 HIV-positive patients, 23 (64%) denied HIV risk factors other than heterosexual contact with persons of unknown HIV status. HIV antibody was associated with GUD (odds radio [OR] = 2.72, 95% confidence interval [CI] = 1.20–6.24), multiple concurrent STDs (OR = 2.51, 95% CI = 1.08–5.81), and a history of crack cocaine use (OR = 2.98, 95% CI = inexact–9.61). Crack use was also associated with GUD (OR = 15.15, 95% CI = 3.27–inexact) and multiple simultaneous STDs (OR = 13.87, 95% CI = 4.62–inexact). In a log-linear model analysis, HIV infection was independently associated with GUD and crack use. HIV infection, genital ulcer disease, and crack cocaine use were more common in women than men.

Conclusions. The association between HIV infection and GUD seen here may be secondary to high-risk sexual behavior, which in turn may be partially attributable to crack cocaine use and drug-related prostitution. The high rate of coinfection with HIV and GUD raises a concern about the increased efficiency of sexual transmission of HIV in this population. (*Am J Public Health.* 1991;81:1576–1579)

Introduction

During the last decade divergent patterns of human immunodeficiency virus (HIV) transmission have been noted in different geographic regions. In North America, HIV transmission has occurred primarily in male homosexuals and intravenous drug users (IVDUs). Studies of heterosexual transmission in the United States have suggested a low efficiency of transmission,[1] while the rate of heterosexual transmission of HIV in Africa appears to be much higher.[2,3] Retrospective and prospective studies suggest that this difference may be related at least in part to the higher prevalence of genital ulcer disease (GUD) in Africa.[3–5] Evidence that GUD may increase the risk of acquiring HIV has also been noted in at least one sexually transmitted disease (STD) clinic population in the United States.[6]

Since 1986 the incidence of primary and secondary syphilis in the United States has been increasing, especially in urban minority heterosexuals.[7] In Brooklyn, this trend has encompassed other etiologies of GUD as well. Between 1986 and 1988, cases of primary and secondary syphilis increased 98% (from 589 to 1165), chancroid cases increased 105% (from 556 to 1140), and herpes simplex genital ulcers increased 78% (from 324 to 577).[8] In the STD clinic at Kings County Hospital in Brooklyn, the number of cases of GUD more than tripled from 1986 to 1988 and by the first half of 1989, 35% of STD clinic visits were for GUD.

The impact of this dramatic rise in GUD on HIV transmission patterns in this population remains to be seen. Because this increase has been relatively recent and the incubation period for symptomatic HIV disease is several years, any resulting shift in the risk factor categories of patients reported with AIDS will not become apparent for several years. In this retrospective study, STD clinic visit diagnoses were correlated with HIV antibody status and HIV risk factor history.

Methods

This study took place at a municipal hospital–based STD clinic serving an urban minority population. Confidential HIV counseling and testing with written informed consent have been offered to STD clinic patients since October 1988. HIV risk behavior history was obtained at the initial visit by trained interviewers using a standard questionnaire and also during the follow-up visit when test results were given to the patient. HIV antibody testing was performed by the New York City Department of Health Retrovirology Laboratory. Positive enzyme immunoassay results were confirmed by Western blot.

STD clinic records of patients who underwent HIV testing from October 1988 through February 1989 were retrospectively reviewed to determine STD diagnoses. Primary and secondary syphilis were diagnosed by darkfield examination

The authors are with the Infectious Diseases Division, Departments of Medicine and of Obstetrics and Gynecology, and the AIDS Prevention Center, Department of Preventive Medicine and Community Health, State University of New York, Health Sciences Center at Brooklyn, New York.

Requests for reprints should be sent to Dr. Keith Chirgwin, Box 56, 450 Clarkson Avenue, Brooklyn, NY 11203.

This paper was submitted to the journal August 20, 1990, and accepted with revisions April 4, 1991.

Editor's Note. See related Editorial by Hinman on page 1557.

or direct monoclonal immunofluorescent staining for *T. pallidum* or by serology (rapid plasmin reagin confirmed by fluorescent treponemal antibody-absorbed) with compatible physical findings. Chancroid was diagnosed by culture or characteristic gram stain morphology. Specimens cultured for *Hemophilus ducreyi* were inoculated onto Mueller-Hinton medium with 5% chocolated horse blood and gonococcal agar base with bovine hemoglobin and fetal calf serum.[9] Herpes simplex was diagnosed by viral culture (specimens inoculated into A549 and human neonatal kidney cell lines) or direct immunofluorescent staining.

Except where noted, Cornfield confidence intervals and Mantel-Haenszel χ^2 tests were used in calculations of statistical significance. Log-linear model analysis was performed using SPSS statistical software (SPSS, Chicago, Ill).

Results

During this 5-month period, 513 new patients were seen and records were available for 194 who consented to HIV testing. HIV risk factor and demographic characteristics of HIV seronegatives and seropositives are listed in Table 1. Twenty-three of 36 HIV seropositives (64%) denied risk factors other than heterosexual contact with persons of unknown HIV status.

Twenty-four (67%) of 36 HIV seropositives presented with GUD (Table 2). Conversely, 24 of 91 (26%) patients with GUD were seropositive for HIV. GUD was associated with HIV seropositivity (odds ratio [OR] = 2.72; 95% confidence interval [CI] = 1.20–6.24; $P < .01$).

Herpes simplex was the only genital ulcer pathogen in 5 (21%) of 24 HIV-positive genital ulcer patients and could conceivably have been increased in this group due to HIV-related immunosuppression. When genital ulcers due to herpes simplex were excluded from analysis, the association between HIV infection and GUD persisted (OR = 2.82; 95% CI = 1.14–7.03; $P < .02$). In six patients who presented with genital ulcers, the etiology of the ulcer was not determined. All but one of these individuals were HIV seropositive.

Two or more concurrent STDs (not including HIV) were diagnosed in 46 patients. Fourteen (30%) of 46 patients with multiple STDs were HIV seropositive compared with 22 (15%) of 148 patients with fewer than two STD diagnoses (OR = 2.51; 95% CI = 1.08–5.81; $P < .02$).

TABLE 1—HIV Seroprevalence by Demographic Characteristics and HIV Risk Factor.

Characteristic	HIV Seropositive n = 36 (%)	HIV Seronegative n = 158 (%)
Median age (years)	28	26
Sex		
Female	26 (72)	79 (50)*
Male	10 (28)	79 (50)
Race		
Black	31 (86)	137 (87)
Hispanic	5 (14)	17 (11)
White	0	4 (2)
HIV risk factor		
Intravenous drug use	9 (25)	4 (3)
Male homosexual	2 (6)	1 (1)
Heterosexual partner known to be HIV infected	2 (6)	2 (1)

*$P < .02$.

TABLE 2.—HIV Seroprevalence by Genital Ulcer Diagnosis

Diagnosis	All STD[a] patients N = 513	No. of patients (%) consenting to HIV testing n = 194	HIV antibody positive (%) n = 36
Primary and secondary syphilis	101	45 (23)	9 (25)
Herpes simplex	48	32 (16)[b]	8 (22)
Chancroid	34	26 (13)[c]	3 (8)
Genital ulcers of uncertain etiology	18	6 (3)	5 (14)
All genital ulcers[d]	183	91 (47)	24 (67)*
Nongenital ulcer diagnosis	330	103 (53)	12 (33)

[a]STD = sexually transmitted disease.
[b]Herpes simplex diagnosed by viral culture and direct immunofluorescent staining in 17, by culture alone in 11, by direct immunofluorescent staining alone in 4.
[c]Chancroid diagnosed by culture and gram stain in 14, culture alone in 4, gram stain alone in 8.
[d]In 18 patients more than one etiologic agent for genital ulcer disease was identified; as a result, number of patients with genital ulcer disease is less than number of genital ulcer diagnoses.
*Odds ratio = 2.72, $P < .01$.

A history of using crack cocaine, a freebase form of cocaine which can be smoked, was noted in 23 patients (18 women, 5 men). Twenty-one (23%) of 91 patients with GUD admitted to crack cocaine use compared with 2 (2%) of 103 patients without GUD (OR = 15.15; 95% CI = 3.27–inexact; $P < .001$). Seventeen (37%) of 46 patients with multiple concurrent STDs admitted to crack cocaine use compared with 6 (4%) of 148 patients with less than two STDs (OR = 13.87; 95% CI = 4.62–inexact; $P < .001$). Two or more etiologies of GUD were diagnosed concurrently in 9 (39%) patients admitting to crack use, compared with 9 (5%) patients without a history of crack use (OR = 11.57; 95% CI = 3.51–38.78; $P < .001$, Fisher's exact).

Six (26%) of the 23 patients with a history of crack use were seropositive for HIV (4 women, 2 men). All six patients denied usual HIV risk factors including intravenous (IV) drug use. Five of these six patients presented with genital ulcers, and the sixth with an oral ulcer. The association between HIV infection and crack use was not statistically significant unless patients with recognized risk factors for HIV (e.g., a history of IV drug use, homosexuality, or a sexual partner known to be infected with HIV) were excluded from analysis. When patients with recognized HIV risk factors were excluded, the HIV seroprevalence rates for patients with and without a history of crack use were 26% (6 of 23) and 11% (16 of 151) (OR = 2.98; 95% CI = inexact–9.61; $P < .05$, Fisher's exact). Similarly, when those with recognized HIV risk factors were excluded from analysis there were increases in the association of HIV

TABLE 3—HIV Infection, History of Crack Cocaine Use, Genital Ulcer Disease, and Multiple Concurrent STDs in STD Clinic Patients without Recognized HIV Risk Factors

HIV Serostatus	Crack Cocaine Use	No. of Patients	No. (%) with GUD	No. (%) with >1 STD
Positive	Yes	6	5 (83)	4 (67)
Positive	No	17	11 (65)	6 (35)
Negative	Yes	16	15 (94)	13 (81)
Negative	No	135	48 (36)	17 (13)
Totals		174	79 (45)	40 (23)

Note. GUD = genital ulcer disease; STD = sexually transmitted disease.

with GUD (OR = 3.68; 95% CI = 1.23–11.29; $P < .01$) and with multiple concurrent STDs (OR = 3.10; 95% CI = 1.13–8.49; $P < .02$). Log-linear model analyses were performed with the following variables: HIV infection, GUD, multiple concurrent STDs, and crack cocaine use (Table 3). Those with recognized HIV risk factors were excluded from analysis. In the model, HIV infection remained independently associated with GUD and crack cocaine use. The association between HIV infection and multiple concurrent STDs, however, depended on the association of each with crack use.

The rate of HIV infection in women was more than double that in men (25% vs 11%: OR = 2.60; 95% CI = 1.11–6.22; $P < .02$). Women were more likely than men to have GUD (OR = 2.28; 95% CI = 1.23–4.26; $P < .006$), multiple concurrent STDs (OR = 2.07; 95% CI = 0.98–4.40; $P < .04$), and history of crack use (OR = 3.48; 95% CI = 1.15–inexact; $P < .02$).

Discussion

In this inner-city, predominantly heterosexual STD clinic population the rate of HIV infection among clients tested was 19%. The majority of HIV-infected individuals denied usual risk factors for HIV such as IV drug use, homosexuality, bisexuality, or having sexual partners known to be infected with HIV. HIV infection was associated with female sex, crack cocaine use, GUD, and multiple concurrent STDs. With log-linear regression analysis, HIV infection remained independently associated with GUD and crack cocaine use.

A similar trend of increased HIV seropositivity in those presenting with GUD in New York City has been described by others.[10] In a city-wide study involving approximately 8600 STD clinic patients in 1988, 13% of those presenting with genital ulcers were seropositive for HIV compared with an 8% overall HIV seropositivity rate.[10]

The time interval to seroconversion for HIV exceeds the usual incubation period for most GUD and therefore HIV infection probably preceded the presenting genital ulcer in most of these individuals. One possibility is that immunosuppressed HIV-infected source partners are at increased risk of developing persistent or recurrent genital ulcers. If this is true, someone who acquires GUD may be more likely than someone who acquires another STD to have been exposed to an immunosuppressed partner with HIV infection. Since HIV transmissibility appears to increase with immunosuppression, such source partners may be more likely to have transmitted HIV even prior to the development of GUD. Another possibility is that genital ulcers may be a direct manifestation of HIV disease.[11] Finally, the explanation for the association between HIV infection and GUD seen in these patients may be that both are dependent covariables of sexual behavior that occurs in milieus where both HIV infection and genital ulcers are common. One such setting may be "crack houses," retail distribution points for crack cocaine where drug-related prostitution is reported to be common.

The rising incidence of syphilis in urban minority populations has been ascribed to a variety of factors, including funding cutbacks in STD control programs, increasing use of spectinomycin (which is not active against incubating syphilis) for the treatment of gonorrhea in areas where beta-lactamase–producing gonorrhea are prevalent, and the recent epidemic of crack cocaine use.[12] Crack cocaine use may be associated with high-risk behavior such as decreased selectivity in sexual partner choice and increased participation in unsafe sexual practices.[12–15] Drug-related prostitution, i.e., sex in exchange for drugs or for money to purchase drugs, may play a critical role in this association.[13,14] In one recent study the association of syphilis with cocaine use persisted after adjusting for number of sexual partners and prostitution, suggesting the possibility of an increase in unsafe sexual behavior independent of prostitution or another factor.[15] In this study, crack use was strongly associated with multiple simultaneous STDs, which may be a reflection of multiple partner sexual behavior with greater opportunity for consequent exposure to multiple STDs, including GUD and HIV.

The association of GUD with HIV infection was more pronounced in women than men. This sex-specific association might be attributable to drug-related prostitution, likely a more prevalent behavior in women than men. If female addicts engage in sexual behavior in the context of crack use more frequently than male addicts, then events that depend on such high-risk sexual behavior, such as HIV infection and genital ulcer disease, will also be more common in women.

Polydrug abuse involving both crack cocaine and IV drugs is a potentially confounding variable that might account for the association of HIV infection and crack. Arguing against this proposition is the observation that the association of HIV infection with crack use increased when patients with recognized HIV risk factors (primarily IV drug users) were excluded from analysis. Furthermore, the association of HIV with GUD and multiple concurrent STDs also increased when patients with recognized risk factors for HIV were excluded from analysis. These observations are consistent with the hypothesis that high-risk sexual behavior may have accounted for HIV infection in these heterosexuals without recognized HIV risk factors.

Since voluntary HIV testing was employed in this study, it is possible that selection bias may have influenced these results, although the magnitude and direction of such bias is difficult to ascertain. It is possible that testing was encouraged more aggressively for patients with recognized risk factors for HIV, including those with multiple heterosexual partners and multiple episodes of STDs. On the other hand, many individuals who perceived themselves at high risk, particularly IVDUs, may have been less likely to consent to testing. The net effect of these biases may have been to underrepresent

IVDUs and overrepresent those with heterosexual risk factors who did not perceive themselves at risk. Despite this limitation, these data may provide important clues to the direction of the HIV epidemic in similar urban minority populations. In New York City, the estimated number of IVDUs has remained essentially stable over the last decade, while the estimated number of non-IV drug users, especially of crack cocaine, has continued to rise.[16] If these trends continue, sexual risk behavior associated with crack use may become the dominant mode of HIV transmission in this population.

Crack use history may also have been biased due to the previously reported association between crack use and GUD.[12] It is possible that a history of crack use was more consistently elicited and recorded in the clinic charts of patients presenting with genital ulcers than those with other STDs. This bias by itself would not account for the association seen with HIV and multiple concurrent STDs because in most cases these diagnoses were based on laboratory results that became available after the patient had left the clinic.

The frequent coexistence of HIV with GUD in this population is of concern for several reasons. Genital ulcers may enhance the efficiency of HIV transmission by several mechanisms. The isolation of HIV from the genital ulcers of HIV-infected individuals[17] raises the possibility that genital ulcers may increase the infectiousness of index partners by resulting in increased shedding of HIV. In addition, susceptibility of the contact who acquires GUD from the index partner may be increased. This might occur by disruption of the genital epithelium during ulcer formation as well as by recruitment of HIV-susceptible activated immune cells to the genital tract.[18,19] Thus, the high prevalence of both HIV and GUD in this population may create conditions similar to those in Africa where heterosexual transmission of HIV appears to be common.

The rising incidence of GUD in this population suggests that current efforts to promote reduction in risk behavior are insufficient. Such risk behavior reduction will play an essential role in controlling the parallel epidemics of GUD and HIV. The future of the HIV epidemic in this country may depend largely on the effectiveness of interventions aimed at similar urban minority populations where GUD and HIV infection are on the rise. Increased efforts should focus not only on STD control and risk behavior reduction, but also on prevention and treatment of non-IV drug use associated with sexual risk behavior. □

Acknowledgments

This research was supported in part by Health Research, Inc., grant 815-3335, and the New York State Department of Health AIDS Institute through its prime contract from the US Public Health Service (U62CCU20206104).

We are indebted to Ramona Vitkevich for providing New York City Department of Health STD data and Paula Freedman, Arnaud Bastien, Yvette Norat, Joseph Anarella, and Janice Handte for their assistance.

References

1. Padian N, Wiley J, Winkelstein W. Male-to-female transmission of human immunodeficiency virus (HIV): current results, infectivity rates, and San Francisco population seroprevalence estimates. Presented at the Third International Conference on AIDS; June 1–5, 1987; Washington, DC. Abstract.
2. Carael M, Van De Perre PH, Lepage PH, et al. Human immunodeficiency virus transmission among heterosexual couples in Central Africa. *AIDS.* 1988;2:201–205.
3. Cameron DW, Simonsen JN, D'Costa LJ, et al. Female to male transmission of human immunodeficiency virus type 1: risk factors for seroconversion in men. *Lancet.* 1989;2:403–407.
4. Kreiss JK, Koech D, Plummer FA, et al. AIDS virus infection in Nairobi prostitutes: spread of the epidemic to East Africa. *N Engl J Med.* 1986;314:414–418.
5. Greenblatt RM, Lukehart SA, Plummer FA, et al. Genital ulceration as a risk factor for human immunodeficiency virus infection. *AIDS.* 1988;2:47–50.
6. Quinn TC, Glasser D, Cannon RO. Human immunodeficiency virus infection among patients attending clinics for sexually transmitted diseases. *N Engl J Med.* 1988; 318:197–203.
7. Centers for Disease Control. Increases in primary and secondary syphilis—United States. *MMWR.* 1987;36:393–397.
8. New York City Department of Health Bureau of STD Control. *Annual Report 1988* New York, NY.
9. Nsamze H, Plummer F, Maggwa ABN, et al. Comparison of media for primary isolation of *Hemophilus ducreyi. Sex Trans Dis.* 1984;11:6–9.
10. New York City Department of Health. Genital ulcer disease in New York City. *City Health Information.* 1989;8:1–3.
11. Covino JM, McCormack WM. Vulvar ulcers of unknown etiology in an HIV-infected woman: response to treatment with zidovudine. *Am J Obstet Gynecol.* 1990;163:116–118.
12. Centers for Disease Control. Continuing increases in infectious syphilis—United States. *MMWR.* 1988;37:35–38.
13. Centers for Disease Control. Congenital syphilis—New York City, 1986–1988. *MMWR.* 1989;38:825–829.
14. Scribner R, Cohen D. The streets of Babylon: syphilis and sex for drugs in Los Angeles County. In: Program and abstracts of the 117th Annual Meeting of the American Public Health Association; Oct. 22–26, 1989; 18. Washington, DC. Abstract.
15. Rolfs RT, Goldberg M, Sharrar RG. Risk factors for syphilis: cocaine use and prostitution. *Am J Public Health.* 1990;80:853–856.
16. Frank B, Hopkins W, Lipton DS. *Current Drug Use Trends in New York City, June 1989.* New York, NY: New York State Division of Substance Abuse; 1989:1–10.
17. Kreiss JK, Coombs R, Plummer F, et al. Isolation of human immunodeficiency virus from genital ulcers in Nairobi prostitutes. *J Infect Dis.* 1989;160:380–384.
18. Lukehart SA, Baker-Zander SA, Lloyd RM, Sell S. Characterization of lymphocyte responsiveness in early experimental syphilis. *J Immunol.* 1980;134:461–467.
19. Sheldon WH, Heyman A. Studies on chancroid, I: Observations on the histology with an evaluation of biopsy as a diagnostic procedure. *Am J Pathol.* 1985;22:415–425.

Reduction of High-Risk Sexual Behavior among Heterosexuals Undergoing HIV Antibody Testing: A Randomized Clinical Trial

Neil S. Wenger, MD, MPH, Lawrence S. Linn, PhD, Marsha Epstein, MD, MPH, and Martin F. Shapiro, MD, PhD*

ABSTRACT

Background. We evaluated the effect of HIV antibody testing on sexual behavior and communication with sexual partners about AIDS risk among heterosexual adults at a clinic for sexually transmitted diseases.

Methods. We randomized 186 subjects to receive either AIDS education alone (the control group) or AIDS education, an HIV antibody test, and the test results (the intervention group). These subjects were then followed up 8 weeks later.

Results. At follow-up, mean number of sexual partners decreased, but not differently between groups. However, compared with controls, HIV antibody test intervention subjects, all of whom tested negative, questioned their most recent sexual partner more about HIV antibody status ($P < 0.01$), worried more about getting AIDS ($P < 0.03$), and tended to use a condom more often with their last sexual partner ($P = 0.05$): 40% of intervention subjects vs 20% of controls used condoms, avoided genital intercourse, or knew their last partner had a negative HIV antibody test ($P < 0.005$).

Conclusion. HIV antibody testing combined with AIDS education increases concern about HIV and, at least in the short term, may promote safer sexual behaviors. Additional strategies will be necessary if behaviors risky for HIV transmission are to be further reduced. (*Am J Public Health.* 1991;81:1580–1585)

Introduction

Antibody testing for human immunodeficiency virus (HIV) has been proposed as part of the effort to control transmission of acquired immunodeficiency syndrome (AIDS).[1,2] Such testing is now routine for military personnel, federal prisoners, immigrants, and, in some states, those applying for marriage licenses.[3] Voluntary HIV antibody testing also is becoming more commonplace in special testing centers, medical practices, and clinics. The question of who should be tested has been debated by physicians,[4] ethicists,[5] activist groups, and presidential candidates,[6] but it remains controversial. Although HIV testing has become widespread, we are only beginning to understand its effects.[7]

Descriptive studies of high-risk individuals (homosexual men and intravenous [IV] drug users) requesting HIV testing and receiving positive test results have demonstrated some decrease in sexual activity among those tested.[8] However, almost no data have been published on the impact of testing on lower-risk populations, such as heterosexual adults sexually active with multiple partners, the overwhelming majority of whom will have negative test results. Yet such information is essential in developing public policy. A negative result on an HIV antibody test could be used either to justify resumption of a high level of sexual activity or to stimulate more cautious behavior.

No randomized trials of HIV testing have been reported to date. Thus, to evaluate the impact of HIV antibody testing on sexually active heterosexual adults, we conducted a randomized trial in a Los Angeles clinic for sexually transmitted disease (STD).

Methods

Consecutive individuals attending an urban STD clinic between January and March 1988 were approached to participate in a randomized trial of HIV antibody testing. The clinic did not offer HIV antibody testing at the time of this study.

Of 724 consecutive patients, 224 were judged ineligible according to preset criteria: they had been previously approached about study participation (21%) or they were younger than 18 years of age (14%), unable to speak English (46%), of homosexual or bisexual orientation (4%), unavailable due to constraints of the clinic (i.e., sequestered by peace officers or primarily participating in other clinic areas [10%]), unable to give informed consent (4%), or unable to give a follow-up address (1%). Of the 500 eligible patients, 259 (52%) were interested in receiving free, confidential HIV antibody testing and 256 (51%) were willing to participate in the randomized trial. Those attending the clinic, those eligible for the study, and those participating were similar in gender and age; there were more Blacks in the study sample than among clinic attendees (84% vs 72%) due to the exclusion of non-English speakers (all of whom were Hispanic).

The authors are with the Division of General Internal Medicine and Health Services Research, Department of Medicine, University of California, Los Angeles, and the Los Angeles County Department of Health Services.

*Dr. Lawrence S. Linn died since this study was completed.

Requests for reprints should be sent to Neil S. Wenger, MD, Department of Medicine, UCLA School of Medicine, B-564 Factor Building, Los Angeles, CA 90024-1685.

This paper was submitted to the journal December 4, 1990, and accepted with revisions July 1, 1991.

Upon entering the clinic, patients were informed of the possibility of receiving free, confidential HIV antibody tests along with the routine clinic blood draw for syphilis serology. The nature of the randomized trial was explained to eligible patients who wished to receive the HIV antibody test. Signed consent was obtained for the study and for HIV antibody testing.

Subjects then completed a self-administered questionnaire that included items to elicit (1) demographic information; (2) knowledge about AIDS (six true/false questions); (3) mental health (a five-item global mood inventory);[9] (4) worry about general health (a four-item "health/worry" index);[10] and (5) information about their last sexual partner, including length of relationship, perceived sexual experience, and chance of carrying HIV. Communication with sexual partners was evaluated by inquiring as to whether subjects had *asked* either partner in the past month or their most recent partner (1) whether that partner had been tested for antibody to HIV and, if so, what the result was; (2) whether that partner had used IV drugs; and (3) how many previous sexual partners that partner had. Subjects' sexual behavior was measured as (1) the number of sexual partners in the past month, (2) the number of episodes of vaginal and anal intercourse and of oral sex with and without a condom in the past month with all sexual partners, and (3) the number of sexual acts with their most recent partner.

After completing the questionnaire, all subjects participated in an educational module in which they (1) received a written pamphlet that explicitly discussed safer and unsafe sexual acts and explained condom use;[11] (2) watched a 15-minute videotape that discussed behavior risky for contracting AIDS and promoted condom use and discussion of risk with sex partners;[12] and (3) participated in a 10-minute, one-on-one counseling session with a physician blinded to randomization status. The counseling focused on assessing personal risk, discussing the elements of the HIV test, and answering any questions about AIDS or HIV testing. All aspects of usual HIV pretest and posttest counseling were covered with every subject.

After completing the educational module, subjects were called for the usual clinic blood draw by the clinic nurse, who opened the sealed randomization envelope. Subjects randomized to HIV testing (the intervention group) were informed of that fact and had an extra tube of blood drawn. Subjects randomized not to receive the test (the control group) were offered a list of locations for free, anonymous HIV antibody testing.

Serum samples underwent enzyme-linked immunosorbent assay (ELISA) for HIV. Repeat positive results on ELISA tests were submitted for confirmatory western blot analyses. Negative test results were revealed to subjects, in person or by telephone, approximately 2 weeks after study entry and were accompanied by the same risk reduction message that all intervention and control subjects received during the pretest counseling. Subjects testing positive received face-to-face, in-depth counseling with analysis of risk factors. Long-term medical and psychological follow-up was arranged.

Follow-up questionnaires were mailed to all subjects 8 weeks after study entry. In addition to the names, addresses, and telephone numbers they had provided, subjects were traced by using public health and social service resources, canvassing nearby hospitals and penal institutions, and monitoring the clinic. In addition to the questions on the first questionnaire, subjects were asked whether they were worried about getting AIDS, whether they had been HIV tested, and what their responses were to testing. If the questionnaire was not returned after repeat mailings and telephone calls, a trained research assistant administered it over the telephone. The number of responses obtained by telephone did not differ between the two groups.

Of the 256 individuals consenting to participate, 125 were randomized to the intervention group and 131 to the control group. Every effort was made to follow up each subject. However, follow-up efforts revealed that 22 intervention subjects (18%) and 23 control subjects (18%) had provided untruthful names, addresses, and telephone numbers when they entered into the study—before randomization. These subjects were thus excluded from the study. The excluded subjects in the two groups were similar with regard to demographic characteristics, communication with partners about AIDS, and sexual behavior (Table 1).

Of the 103 remaining intervention subjects, 11 (11%) could not be followed up (1 left the clinic before blood could be drawn, 2 could not be reached to receive test results, 1 refused to be informed of his [positive] result, and 7 could not be located or refused to complete follow-up questionnaires). Similarly, of the remaining 108 control subjects, 14 (13%) did not complete follow-up questionnaires. Those subjects not followed up were similar at baseline between the two groups except for small differences in age and knowledge about AIDS. Compared with those who were followed up, they were somewhat less likely to be Black (Table 1). Overall, the 70 subjects who were not followed up in this trial did not differ significantly from those who were followed up in terms of demographics (except race), knowledge about AIDS, number of sexual partners last month, or previous episodes of STD. The remaining 92 subjects in the intervention group and 94 subjects in the control group make up the longitudinal sample for this study.

Three of the 124 subjects receiving an HIV antibody test through this study tested positive for HIV on ELISA. All were confirmed on western blot analysis. At follow-up counseling, one revealed that he had denied his homosexual contact to enter the study. The second was a woman with two children and an IV drug–using ex-husband; she received follow-up care but refused to complete the follow-up questionnaire. The third was a male who declined to receive his test result, fearing that he might become depressed if he were positive; he also declined to complete the follow-up questionnaire. Thus, all subjects who were followed up in the intervention group received a negative HIV antibody test result approximately 6 weeks prior to completing the follow-up questionnaire. Among the 94 control group subjects completing the follow-up survey, 10 (11%) obtained HIV antibody tests elsewhere. (Nine reported a negative result and one never obtained the result.)

Analyses were performed only on prespecified hypotheses using an intention-to-treat protocol in which subjects were analyzed in their original randomized groups, regardless of whether persons in the control group obtained the HIV test from an outside source. Baseline differences between intervention and control groups for continuous variables were assessed with t tests or Wilcoxon tests, depending on the variable's underlying distribution; differences for dichotomous variables were assessed with chi-square tests. Repeated measures analysis of variance for continuous variables (using SAS Proc GLM) and log-linear models for dichotomous variables (using SAS Proc CATMOD) were used to evaluate the effects of group differences, changes over time, and interactions.

TABLE 1—Baseline Comparison of the Study Sample to Those Providing No Follow-up Information and Those Lost to Follow-up

	No Follow-up Information		Lost to Follow-up		
	Randomized to Intervention Group	Randomized to Control Group	Randomized to Intervention Group	Randomized to Control Group	Followed-up
	(n = 22)	(n = 23)	(n = 11)	(n = 14)	(n = 186)
Age (mean ± standard deviation [SD])	25 ± 5	27 ± 5	24 ± 8*	31 ± 9*	27 ± 8
Male (%)	68	61	73	64	67
Black, %**	82	74	73	71	88
Unmarried, %	96	100	91	93	86
Education, median in years	13	13	12	12	13
Employed, %	68	52	55	50	57
Number of sexual partners:					
In the last month, mean ± SD	2.9 ± 4.3	3.2 ± 2.2	1.6 ± 1.1	1.6 ± 1.5	1.8 ± 1.5
Lifetime, median	20	39	18	18	14
Number of episodes of sexually transmitted diseases, mean ± SD	3.1 ± 2.8	3.6 ± 3.2	2.8 ± 3.1	2.1 ± 2.2	2.7 ± 2.6
AIDS knowledge index score, mean number correct of six questions ± SD	5.1 ± 1.0	4.9 ± 1.1	4.3 ± 1.0*	5.4 ± 0.9*	5.0 ± 1.1
Mental health scale score, mean ± SE	4.5 ± 0.2	4.5 ± 0.2	4.5 ± 0.2	4.1 ± 0.2	4.6 ± 0.1
Health/worry scale score, mean ± SE	4.2 ± 0.2	3.8 ± 0.1	3.7 ± 0.3	4.1 ± 0.2	4.1 ± 0.1
Number of AIDS risk questions asked of last partner, mean ± SD	0.4 ± 0.7	0.8 ± 1.0	0.6 ± 0.7	0.9 ± 1.0	0.9 ± 1.0
Unprotected vaginal or anal intercourse with last partner, %	82	83	100	79	90

*$P < 0.05$ for t-test comparison of means of intervention and control groups.
**$\chi^2 = 4.6$, $P < 0.04$ for comparison between lost to follow-up (n = 70) and followed-up (n = 186) groups.

TABLE 2—Baseline Characteristics of the Intervention and Control Groups

	Intervention Group	Control Group
	(n = 92)	(n = 94)
Demographic Characteristics		
Age, range 18, 66; mean in years ± SD	28 ± 8	27 ± 8
Male, %	71	63
Black, %	93*	82*
Completed high school, %	85	83
Unmarried, %	84	88
Employed, %	55	60
Income, % less than $1000 per month	73	73
Knows person with AIDS, %	5	7
Previous HIV antibody test, %	5	9
Age at first intercourse, range 3, 26; median in years	15	15
Previous episodes of sexually transmitted disease, range 0, 13; mean ± SD	2.6 ± 2.6	2.8 ± 2.6
Sexual Behavior		
Number of sexual partners		
Past 30 days, range 0,15; mean ± SD	1.9 ± 1.7	1.7 ± 1.2
Same month last year, range 0,20; mean ± SD	3.1 ± 3.4	2.6 ± 2.3
Lifetime, range 1, 1000; median	16	13
Vaginal or anal intercourse without a condom with last sexual partner, %	90	89

*$\chi^2 = 4.7$, $P = 0.03$ for difference between groups.

Results

The characteristics of subjects completing the study are shown in Table 2. The mean age was 27 years. Most were male, Black, and high school graduates and were unmarried, employed, and reported an income of less than $1000 per month; 6% reported knowing someone with AIDS, 5% used IV drugs, and 7% had been previously tested for antibody to HIV. Only race differed significantly between the groups ($P = 0.03$). Baseline scores for knowledge about AIDS, mental health, and health/worry were similar.

At baseline, subjects reported a mean of 1.8 sexual partners in the past month compared with 2.9 for the same month the previous year ($P < 0.001$). They reported a median of 14 lifetime sexual partners, age at first sexual intercourse of 15, and a mean of 2.7 previous episodes of STD. The groups did not differ in these baseline values or in the proportion engaging in various sexual acts with all partners over the past month (Table 2).

The two groups did not differ in length of sexual relationship with the last partner, in estimated number of previous sexual partners that this last partner had had, or in estimates of this partner's chance of carrying HIV. Nearly half reported having asked if their last partner had used injected drugs, 14% had asked if that partner had been tested for HIV, and 26% had asked how many previous sexual partners this partner had had. There were no significant differences between groups at study entry in these measures of communication either with the last partner or with all partners in the past month. Re-

garding their last sexual encounter, 88% reported vaginal intercourse without condoms (12% with condoms) and 9% reported anal intercourse without condoms (1% with condoms). The groups did not differ in baseline measures of sexual activity (Table 3).

At follow-up, there were no differences between intervention and control groups or from baseline in measures of AIDS knowledge, mental health, or health/worry (Table 4). However, when specifically asked at follow-up if they were concerned about AIDS, intervention subjects expressed greater concern than controls ($P < 0.03$). Comparing those who were more worried with those who were not, intervention group subjects were significantly more likely to be *more* worried about getting AIDS than they had been 1 month previously (45% vs 26%, $P = <0.01$).

Intervention subjects were more likely than control subjects to report at follow-up that they had asked their last sexual partner about their risk of carrying HIV (Table 5). Inquiries about that partner's HIV status rose from 13% to 41% in the intervention group, significantly more than the 15% to 24% rise in the control group ($P < 0.01$, difference = 18%, 95% confidence interval [CI]: 7.4%, 29%). This difference was due to two effects: 29% of intervention subjects *had not* asked about HIV status at baseline but did so at follow-up, compared with 17% in the control group ($P = .01$), and only 1% of intervention subjects who *had* asked about HIV status at baseline "backslid" and failed to do so at follow-up, compared with 9% of control subjects ($P = .01$).

The proportion of subjects who reported having asked their most recent sexual partner about that person's number of prior sexual partners increased in both groups: from 23% to 57% in the intervention group and from 29% to 53% in the control group. Fifty percent of intervention subjects asked about IV drug use at follow-up compared with 34% of control subjects; however, differences between groups were not significant (Table 5).

Overall, 51% of intervention subjects and 32% of control subjects asked their last sexual partner more questions at follow-up than at baseline about their risk of carrying HIV (difference = 19%, 95% CI: 5%, 33%).

As might be expected after a visit to an STD clinic, the last sexual partner was a first-time partner for only 14% of subjects at follow-up (vs 24% at baseline, $P < 0.05$). There were no differences at follow-up between groups in the percentage for whom the most recent sex partner was a new partner, in the length of sexual relationship with that sexual partner, in the subject's perception of that partner's prior sexual experience, or in the subject's perception of that partner's risk of carrying HIV.

Subjects in both groups decreased their total number of sexual partners per month from 1.9 and 1.7 partners in the intervention and control groups, respectively, at study entry to 1.4 and 1.3 partners, respectively, at follow-up ($P < 0.01$), though the groups did not differ (Table 5).

Although the decline in the number of partners with whom the subjects had vaginal, oral, or anal sex was similar between the two groups, 27% of intervention subjects avoided vaginal or anal intercourse *without a condom* with their most recent sexual partner compared with 13% of control subjects. This approached statistical significance ($P = 0.05$) for a difference between groups. Analyzed separately, the

TABLE 3—Baseline Sexual Behavior with Last Sexual Partner

	Intervention Group (n = 92)	Control Group (n = 94)
Last encounter was first sexual experience with this partner (%)	20	28
Estimate of partner's lifetime number of previous sexual partners, range 1, 15; median	6.1	7.5
Percentage estimating partner's risk of carrying HIV as being less than one in a million	39	46
Sexual activity with last partner (%)		
Vaginal intercourse		
Without condom	89	88
With condom	13	12
Oral-vaginal sex	24	23
Oral-penile sex		
Without condom	33	29
With condom	1	2
Anal intercourse		
Without condom	9	9
With condom	1	0

TABLE 4—AIDS Knowledge, Mental Health, Health/Worry and AIDS Worry Baseline vs Follow-up

	Intervention Group (n = 92)	Control Group (n = 94)
AIDS Knowledge Index Score, mean correct of six questions ± SD		
Baseline	5.0 ± 1.1	5.0 ± 1.0
Follow-up	4.8 ± 1.1	4.9 ± 1.0
Mental health scale score mean ± SE		
Baseline	4.5 ± 0.1	4.8 ± 0.1
Follow-up	4.4 ± 0.1	4.5 ± 0.1
Health/worry scale score, mean ± SE		
Baseline	4.1 ± 0.1	4.1 ± 0.1
Follow-up	4.1 ± 0.1	3.9 ± 0.1
Worry about getting AIDS compared with 1 month earlier (%)*		
More worried	45	26
About the same	37	51
Less worried	18	23

*$\chi^2 = 7.3$, $P < 0.03$ for difference between groups.

TABLE 5—Sexual Behavior and Communication about AIDS Baseline vs Follow-up

	Intervention Group (n = 92)	Control Group (n = 94)
Number of Sexual Partners in the Past 30 Days		
Baseline	1.9 ± 1.7	1.7 ± 1.2
Follow-up	1.4 ± 1.0*	1.3 ± 0.8*
Asked Last Sexual Partner:		
If tested for HIV, %:		
Baseline	13	15
Follow-up	41**	24**
If used injected drugs, %:		
Baseline	47	45
Follow-up	50	34
About number of previous sexual partners, %:		
Baseline	23	29
Follow-up	57	53
Avoided Vaginal or Anal Intercourse without a Condom with Last Sexual Partner, %:		
Baseline	10	11
Follow-up	27***	13***
Avoided Vaginal or Anal Intercourse without a Condom *or* Knew Last Sexual Partner's HIV Test Result Was Negative, %:		
Baseline	13	17
Follow-up	40****	20****

*$P < 0.01$ for change over time.
**$P < 0.01$ for group differences in change over time.
***$P = 0.05$ for group differences in change over time.
****$P < 0.003$ for group differences in change over time.

intervention group avoided intercourse without a condom more at follow-up than at baseline ($P < 0.0001$), whereas the control group did not ($P = 0.34$). Three subjects in the intervention group and none in the control group reported having only oral sex; all other subjects in both groups used condoms for vaginal or anal intercourse (Table 5).

The difference between the groups in the rate of "protected" sexual activity was greater when knowledge of a partner's risk factors was taken into account. Among those having vaginal or anal intercourse without condoms with their last partner, 13 intervention subjects (14%) and 6 control subjects (6%) had asked about this partner's HIV serostatus and had been told the partner was seronegative. Thus, at follow-up, fully 40% (37 of 92) of intervention subjects used a condom, had only oral sex, or stated that they knew their partner's HIV serostatus was negative, whereas only 20% (19 of 94) of control subjects did so ($P < 0.003$, difference = 20%, 95% CI: 7.1%, 33%).

Discussion

Nonrandomized interventions that are intended to change sexual behavior are difficult to evaluate because of the underlying trend toward decreasing rates of activities at risk for transmission of HIV. In this study, all subjects received AIDS education and the components of pre- and posttest counseling, and all were randomized to receive or not receive an HIV antibody test. Compared with individuals in the control group, intervention subjects worried more about getting AIDS (even though they received *negative* HIV test results) and were more likely to have asked partners about AIDS risk factors. HIV testing also was associated with a decrease in unprotected vaginal and anal intercourse. The intervention did not affect our measures of knowledge about AIDS, health/worry, or mental health status.

The behaviors protective for HIV seen among the intervention group in this study are consistent with findings in some observational studies of homosexual men. For example, in an observational study of homosexual men apprised of negative HIV antibody status,[13] van Griensven and colleagues found fewer sexual partners, more communication with partners about HIV serostatus, and increased condom use. Coates et al. found that homosexual men who reported a negative HIV status were more likely to reduce unprotected anal intercourse than those choosing not to be tested.[14,15] Three other observational studies, however, found no substantive behavioral change associated with knowledge of a negative test result in homosexual men.[16–18]

The findings in the current study of greater worry about AIDS and decreased risky sexual behavior among individuals testing negative for HIV are particularly important in light of speculation that a negative test result would adversely affect perceived risk of HIV infection and would be viewed as license to continue "business as usual."[19,20]

While condom use is generally regarded as protective against HIV transmission, some contend that communication with partners about their risk of carrying HIV often is not very useful[21] and might be deleterious if a partner lies.[22] On the other hand, Hearst and Hulley suggest that the sexual partner's probability of HIV infection is "by far the most important" predictor of infection risk; they state that "the best advice we can give our patients is to choose their partners carefully."[23] In this study, intervention subjects increased questioning of partners about HIV testing significantly more than controls did. As with changes in condom use, this may reflect heightened concern about becoming infected with HIV. However, it also may be a result of having been tested; perhaps individuals who know their personal test result are more willing to ask about HIV serostatus.

There is a danger that individuals at STD clinics might fail to understand the "window period" before seroconversion and be incorrectly reassured of negative serostatus in pursuing future unprotected sexual behavior. Although intervention subjects were more likely to use condoms, the majority still had unprotected vaginal or anal intercourse with their last partner. Educational efforts need to stress the meaning of a negative HIV test result, especially during this period after an STD.

The findings of this study need to be interpreted with caution. While intervention subjects significantly decreased risky behavior compared with controls, a large proportion continued to engage in activities at risk for transmission of HIV. The sexual behavior reported in this study is self-reported; subjects may have been untruthful, but it is doubtful that the propensity to lie differed between groups. In addition, 27% of the original sample could not be followed up. Finally, multiple comparisons were performed; however, each was a preplanned analysis, and the fact that 4 of 10 were associated with P values

of .05 or less, all in a similar direction, is more than would have been expected by chance alone.

Furthermore, this study was conducted among a predominantly Black, unmarried, low socioeconomic status population attending an STD clinic. Although this group is of considerable importance because their behaviors may promote the spread of HIV infection among the general heterosexual population,[24–26] findings may not be generalizable to other heterosexuals at lower risk of HIV infection. Half the clinic population was unwilling to receive an HIV test. Indeed, this study compares only the 26% of the STD clinic population that met inclusion criteria, were willing to receive an HIV test and participate in a clinical trial, and were followed up. Results of the trial may apply only to this select group. Lastly, the follow-up period for this study was short. Behavior alterations found at 2 months of follow-up may not have persisted.

This study was carried out as a research project offering free, confidential HIV testing. The fact that subjects accepted testing under such circumstances does not predict the acceptance of testing under other conditions.[27] Perhaps the benefits of testing would be diminished if strict confidentiality could not be assured.[28] Discrimination resulting from reporting of positive HIV test results may produce negative effects not seen in this trial.[29] Confidentiality, sensitivity, and efforts to reduce the stigma of positive HIV test results must be part of effective testing programs.

The findings of this study provide assurance that HIV testing in at-risk heterosexuals is safe: subjects, all of whom tested negative for HIV, did not increase risky sexual behavior and did not suffer adverse psychological consequences of testing. The findings also suggest that, for this heterosexual population, HIV antibody testing is effective in reducing certain risky sexual behaviors, perhaps through the mechanism of increasing the level of worry about acquiring HIV infection. Further studies of testing are needed to assess its long-term effects and to examine its effects in other heterosexual populations with different risk profiles. Finally, although one-time interventions such as that described here may reduce at-risk activities, they do not eliminate them. Additional interventions may well be required to reduce further the risk of HIV transmission. □

Acknowledgments

This research was supported in part by the Robert Wood Johnson Foundation Clinical Scholars Program. However, the views expressed are not necessarily those of the Robert Wood Johnson Foundation.

This paper was presented in part at the American Federation for Clinical Research in Washington, DC, in May 1989.

We are indebted to Dr. Srinivasa Murthy and Carolyn Dong for assistance with the study sample, to Drs. Lee H. Hilborne and Denis O. Rodgerson for HIV antibody testing, to Marlene Nishimoto and Susan Carrasco for technical assistance, and to Dr. Robert H. Brook for invaluable advice.

References

1. Goedert JJ. What is safe sex? *N Engl J Med.* 1987;316:1339–1342.
2. Centers for Disease Control. Public Health Service guidelines for counselling and antibody testing to prevent HIV infection and AIDS. *MMWR.* 1987;36:509–515.
3. Gostin LO. Public health strategies for confronting AIDS. *JAMA.* 1989;261:1621–1630.
4. Experts on AIDS, cite new data, push for testing. *New York Times.* April 24, 1989:1.
5. Bayer R, Levine C, Wolf SM. HIV antibody screening. *JAMA.* 1986;256:1768–1774.
6. Campaign '88: candidates' views. *The Nation's Health.* 1988;May-June:6.
7. Coates TJ, Stall RD, Kegeles SM, Lo B, Morin SF, McKusick L. AIDS antibody testing: will it stop the AIDS epidemic? *Am Psychol.* 1988;11:859–864.
8. Becker MH, Joseph JG. AIDS and behavioral change to reduce risk: a review. *Am J Public Health.* 1988;78:394–410.
9. Steward AL, Hayes RD, Ware JE. The MOS short-form general health survey. *Med Care.* 1988;26:724–735.
10. Brook RH, Ware JE, Davies-Avery A, et al. Overview of adult health status measures fielded in Rand's health insurance experiment. *Med Care.* 1979;17:S1–S131.
11. *Safer Sex: A Guide for Everyone Concerned about AIDS.* Los Angeles, Calif: Los Angeles County Health Department; nd.
12. *Sex, Drugs and AIDS.* New York, NY: ODN Productions; 1987.
13. van Griensven GJP, de Vroome EMM, Tielman RAP, et al. Effect of human immunodeficiency virus (HIV) antibody knowledge on high-risk sexual behavior with steady and nonsteady sexual partners among homosexual men. *Am J Epidemiol.* 1989;129:596–603.
14. Coates TJ, Morin SF, McKusick L. Behavioral consequences of AIDS antibody testing among gay men. *JAMA.* 1987;258:1889. Letter.
15. Coates TJ, Morin S, McKusick L, et al. Long-term consequences of AIDS antibody testing on gay and bisexual men. Presented at the Fourth International Conference on AIDS; June 12–16, 1988; Stockholm, Sweden. Abstract.
16. McCusker J, Stoddard AM, Mayer KH, Zapka J, Morrison C, Sauzman SP. Effects of HIV antibody test knowledge on subsequent sexual behaviors in a cohort of homosexually active men. *Am J Public Health.* 1988;78:462–467.
17. Fife KH, Jones RB, Marrero DG, Katz BP, Serpe RT, Scott J. Behavioral changes among sexually active homosexual men after learning they are negative for HIV antibody. Presented at the Fourth International Conference on AIDS; June 12–16, 1988; Stockholm, Sweden. Abstract.
18. Fox R, Ostrow D, Valdisseri R, Van Raden M, Visscher B, Polk BF. Changes in sexual activities among participants in the multicenter AIDS cohort study. Presented at the Third International Conference on AIDS; June 1–5, 1987; Washington, DC. Abstract.
19. McCombie SC. The cultural impact of the "AIDS" test: the American experience. *Soc Sci Med.* 1987;317:238–241.
20. Fox R, Odaka NJ, Brookmeyer R, Polk BF. Effect of HIV antibody disclosure on subsequent sexual activity in homosexual men. *AIDS.* 1987;1:241–246.
21. Padian NS, Francis DP. Preventing the heterosexual spread of AIDS. *JAMA.* 1988;260:1879. Letter.
22. Cochran SD, Mays VM. Sex, lies, and HIV. *N Engl J Med.* 1990;322:774–775.
23. Hearst N, Hulley SB. Preventing the heterosexual spread of AIDS. *JAMA.* 1988;259:2428–2432.
24. Haverkos HW, Edelman R. The epidemiology of acquired immunodeficiency syndrome among heterosexuals. *JAMA.* 1988;260:1922–1929.
25. Handsfield HH. Heterosexual transmission of human immunodeficiency virus. *JAMA.* 1988;260:1943–1944.
26. May RM, Anderson RM. Transmission dynamics of HIV infection. *Nature.* 1987;326:137–142.
27. Kegeles SM, Catania JA, Coates TJ, Pollack LM, Lo B. Many people who seek anonymous HIV-antibody testing would avoid it under other circumstances. *AIDS.* 1990;4:585–588.
28. Calsyn DA, Freeman G, Saxon AJ, Whittaker S. Addict beliefs about access to HIV test results. Rockville, Md: National Institute on Drug Abuse; 1989(95):417. NIDA Research Monograph.
29. Blendon RJ, Donelan K. Discrimination against people with AIDS: the public's perspective. *N Engl J Med.* 1988;319:1022–1026.

Behavior Changes after Notification of HIV Infection

ABSTRACT

Background. To learn more about how people who did not volunteer for testing react to information about HIV infection, we assessed short-term behavior changes in HIV-positive blood donors.

Methods. Blood donors who were notified at the New York Blood Center that they were HIV positive were asked to participate in a study. A nurse elicited a medical history, performed a limited medical examination, and asked participants to complete a questionnaire that included questions about drug use, sexual behavior, and psychological characteristics. Participants were asked to return in 2 weeks to complete another questionnaire.

Results. Many fewer men and women reported engaging in unsafe sexual behaviors in the 2 weeks preceding the follow-up visit than had reported such behaviors prior to notification. These changes were greater than those other investigators have reported, but about 40% of the participants still reported unsafe sexual activity at the follow-up interview.

Conclusions. To make nonvolunteer screening programs for HIV infection more effective in reducing the spread of HIV infection, we need to learn more about how to help people change their high-risk behaviors. (*Am J Public Health.* 1991;81:1586–1590)

Paul D. Cleary, PhD, Nancy Van Devanter, RN, Theresa F. Rogers, PhD, Eleanor Singer, PhD, Ruth Shipton-Levy, RN, Melanie Steilen, RN, Ann Stuart, RN, Jerry Avorn, MD, and Johanna Pindyck, MD

Introduction

The main goal of blood center screening programs for human immunodeficiency virus (HIV) is to ensure the safety of the blood supply. Another important goal should be to provide information and support to HIV-positive donors to help them reduce the probability that they will infect others.[1,2] The data are inconclusive as to whether screening and notification programs result in positive behavior changes.[3,4] Several researchers have found that HIV-positive individuals showed greater reductions in risk than those who tested negative or who had not been tested,[3,5] but others have found no changes or changes only in certain behaviors.[6–9]

Almost all published reports of the impact of HIV testing and counseling programs are based on data from self-identified homosexual or bisexual men who volunteered for testing. Since many current and proposed screening programs involve people with different risk factors who did not volunteer for testing, it is critical to understand more about how such persons react to information about HIV infection. In this paper, we describe short-term behavior changes in blood donors who were notified that they were HIV-antibody–positive.

Methods

Screening and Notification

Since 1983, the New York Blood Center (NYBC) has asked all donors to indicate on a confidential form whether their blood should be used for transfusions or for studies only.[10] Donors are told to indicate that their blood should be used only for research ("studies only") if they have a risk factor for HIV infection. The NYBC removes all such blood from distribution, regardless of antibody screening results.

Since April 1985, the NYBC has supplemented the confidential self-exclusion program by screening all blood donations for HIV antibodies with enzyme-linked immunosorbent assays (ELISAs). Any sample with two positive ELISA tests and a confirmatory Western blot assay is considered "confirmed positive."

All HIV-positive persons who donated blood between June 1, 1986, and February 28, 1988, and who came to the Center after notification were asked to participate in this study.[1] A letter was mailed to all confirmed HIV-positive donors, asking them to come to the NYBC to discuss a "finding of importance to your health." If two attempts to contact a donor failed, a certified letter was sent explaining the test results and encouraging the donor to come to the Center or to contact a physician for counseling.

A nurse clinician with psychiatric training saw each seropositive donor who

Paul D. Cleary is with the Department of Health Care Policy and Jerry Avorn is with the Department of Social Medicine, both at Harvard Medical School. Nancy Van Devanter is with the Columbia School of Public Health, and Theresa F. Rogers and Eleanor Singer are with Columbia University. Johanna Pindyck is with the AIDS Program and the Transfusion Service at Interfaith Hospital, New York City.

Requests for reprints should be sent to Paul D. Cleary, PhD, Department of Health Care Policy, Harvard Medical School, 25 Shattuck Street, Parcel B—1st Floor, Boston, MA 02115.

This paper was submitted to the journal January 17, 1991, and accepted with revisions July 17, 1991.

came to the Center and explained the meaning of the test results, the potential impact of infection on the donor's health, and ways to prevent the spread of the virus to others. Donors were also given written materials about HIV. The nurses gave the donors counseling and emotional support, answered questions, and assessed the donors' ability to cope with the information.

Sociodemographic and Medical Information

For all blood donors, the NYBC records age, gender, date of donation, and whether the donor used the confidential self-exclusion option. For each study participant, the nurse elicited a complete medical history, including information about possible HIV risk factors, and performed a limited medical examination. A tube of blood was drawn for use in a verification Western blot test and other blood tests. At the end of the first visit, the nurse recorded information about any HIV risk factors that the donor reported.

Questionnaire

After the medical exam and interview, the nurse asked each study participant to complete a questionnaire, in private, that asked about their use of intravenous drugs and whether they had engaged in numerous specific sexual activities, including vaginal sex and various types of anal and oral sex, in the previous 2 weeks. The questionnaire also asked men whether they used a condom always, almost always, sometimes, rarely, or never. Using the same categories, it asked women how frequently a partner used a condom with them. The baseline questionnaire asked participants how many sexual partners of the same sex and how many of the opposite sex they had had in the past month and in the past year, and whether they thought of themselves as exclusively heterosexual (straight), predominately heterosexual, equally homosexual and heterosexual (bisexual), predominately homosexual, or exclusively homosexual (gay).

The questionnaire also included questions about psychological characteristics, such as self-esteem, perceived locus of control, and optimism; sexual and social functioning; social support; coping strategies; health concerns; and depressive symptoms.[11] At the follow-up visit, participants were asked to complete another questionnaire, which contained questions about sexual behavior in the preceding 2 weeks.

Analyses

If information from the structured interview, questionnaire, or nurse clinician summary indicated the presence of a risk factor, we assumed that risk factor to be present. For example, if a male donor reported on the questionnaire that he had had male sexual partners in the previous year, he was classified as having male sexual contact as a risk factor, even if he reported contradictory information during the interview.

We considered participants to have engaged in unsafe sex if they reported that they had engaged in vaginal or anal intercourse and had not always used a condom, that they had engaged in oral sex without a condom, or that they had been a recipient or giver of oral-anal sex.

We compared subgroups of seropositive donors using chi-square statistics for categorical variables and *t* tests for continuous variables. We estimated the associations between donor characteristics and sexual behavior at the 2-week follow-up visit with Pearson product-moment correlations for continuous variables, and point-biserial correlations or fourfold point correlations for dichotomous variables.[12] To estimate the independent effect of donor characteristics on subsequent behavior, we estimated a logistic regression model. We used a forward stepwise selection procedure to select variables that were statistically significant predictors ($P < .05$).

Results

Nonparticipants

During the study period, there were 708 confirmed HIV-antibody–positive donors at the NYBC. Letters to 79 donors (11%) were undeliverable or unclaimed. Another 90 donors (13%) did not respond to the letters from the Blood Center and were notified of their test results by certified mail. Of the 539 donors who contacted the Blood Center, 5 were informed of their results by telephone; 140 agreed to come to the site regularly only to have results of blood tests monitored; 1 was enrolled in another research protocol; 2 were autologous transfusions and were excluded. Two hundred seventy-one patients enrolled in the study, representing 50% of the donors who contacted the study team and 38% of detected seropositive donors. Two hundred and sixty-one of the participants completed an initial interview and 243 completed a baseline questionnaire. Eight study participants completed neither a questionnaire nor a medical interview at the initial visit.

Participating donors were somewhat more likely than nonparticipants to be female (22% vs 18%) and to have said that their blood should be used for studies only (75% vs 69%), but these differences were not statistically significant. The average ages of the two groups were almost identical.

One hundred ninety-six study participants (72%) returned for the 2-week follow-up visit. The median follow-up time was 16 days (mean = 23 days). Those who did not return for the follow-up visit were not significantly different from those who did with respect to CD4 count, age, gender, ethnicity, education, marital status, risk factors, self-exclusion status, period of donation, any of the psychological characteristics measured, or most of the coping strategies measured. However, those who did not return were more likely to be cautious about potential casual routes of transmission (e.g., letting others use towels or silverware that he or she had used) ($P < .05$), less likely to talk with friends to cope with their fears and worries ($P < .01$), and less likely to make themselves feel better by letting out anger ($P < .05$).

Sociodemographic Characteristics

The study population was predominately male, single, young, white, and highly educated, reflecting the characteristics of blood donors in general, although it was very heterogeneous (Table 1).

Risk Factors

The risk factors identified for both men (Table 2) and women (Table 3) are similar to those for the general population of HIV-infected persons. The most commonly reported risk factor for men was sexual contact with another man. Sexual contact with IV drug users and use of intravenous drugs were also reported frequently. A surprising finding was that many men and women who donated for transfusion apparently were not aware of the risk factors for infection.

Reported Behavior Prior to Notification

Most men (63.1%) reported having had a male sexual partner in the previous year, but more than a third had not; 40% reported having had a male sexual partner in the previous month (Table 4). More than 40% of the men reported having had one or more female sexual partners in the

TABLE 1—Sociodemographic Characteristics and Self-Exclusion Status of Seropositive Study Participants

	% Participants
Reason Donated (n = 271)	
For transfusion	69.0
For studies only	31.0
Gender (n = 271)	
Male	77.9
Female	22.1
Age (n = 271)	
18–24	21.0
25–29	31.7
30–34	17.3
35–44	19.9
45–54	7.4
55+	2.6
Education (n = 260)	
Did not complete high school	10.4
High school graduate	21.9
Some college or post–high school training	46.5
College graduate	12.7
More than college	8.5
Marital Status (n = 260)	
Never married	63.1
Married	20.8
Separated or divorced	14.6
Widowed	1.5
Race/Ethnicity (n = 259)	
White, non-Hispanic	44.8
White, Hispanic	19.7
Black, non-Hispanic	30.5
Black, Hispanic	3.5
Other	1.5

TABLE 2—Risk Factors for HIV Infection Reported by Male Seropositive Study Participants

	% Participants Reporting	
	Donated For Transfusion (n = 138)	Donated for Studies Only (n = 73)
Sexual contact with a male since 1977	59.4	93.2
Used IV drugs	19.6	16.4
Sex with an IV drug user since 1977	38.4	35.6
None of above risk factors identified during initial contact	18.1	4.1

TABLE 3—Risk Factors for HIV Infection Reported by Female Seropositive Study Participants

	Participants Reporting	
	Donated For Transfusion (n = 49)	Donated for Studies Only (n = 11)
Sexual contact with a homosexual male since 1977	20.4%	7
Used IV drugs	12.2%	4
Sex with an IV drug user since 1977	65.3%	9
None of above risk factors identified during initial contact	32.7%	1

previous year, and almost a quarter of them reported having had a female sexual partner in the previous month.

Women were much more likely than men to report having had a sexual contact with the opposite sex in the previous year (66.7%), and the vast majority (78.7%) reported no sexual contact with another woman in the previous year. About 40% of both men and women reported that they always used condoms with their sexual partners.

More than 80% of both men and women reported that they had been sexually active in the week prior to notification and more than half had engaged in unsafe sexual practices (Table 5). The rates of sexual activity were comparable for men and women, although men were more likely to have engaged in unsafe sex. Fifty-four men and 9 women who reported that they had had no sex partners in the past month reported that they had engaged in specific sexual activities with others. This discrepancy may have been due to the fact that the questions about partners asked "How many *different* partners of the same (opposite) sex have you had?" (italics added). Some respondents may have answered zero if they had had only a single partner.

Reported behaviors also were not always consistent with self-identified risk factors. Thirty-four percent of the men who said that they thought ofthemselves as exclusively heterosexual reported having had sex with another man.

Behavior Reported at the Follow-Up Visit

At the follow-up interview, almost two thirds of both men and women reported having been sexually active in the previous week, and more than one third of both men and women reported engaging in unsafe sex. To test for the possibility that there had been differences in the frequency of unsafe behavior in different time periods, we compared the proportion of subjects who reported unsafe behaviors in three approximately equal periods of the study and found no consistent differences.

The strongest correlate (not presented here) of whether respondents had engaged in unsafe sex in the period between the two interviews was whether they had engaged in unsafe sex in the week prior to the initial (notification) interview. Only 15.4% of the respondents who reported having had no unsafe sex in the week prior to notification reported having unsafe sex in the period preceding the follow-up interview, but 54% of those who reported having had unsafe sex in the week prior to the initial interview also reported having had unsafe sex before the follow-up interview. Persons who reported that they had had unsafe sex in the period before the follow-up interview also tended to be younger, said at the time of donation that their blood should be used for studies only, reported having had more sexual partners than others at the initial interview, were less concerned about casual transmission, and were less likely to report engaging in a range of health-promoting activities.

In a logistic regression model (Table 6), the independent significant predictors of unsafe sex were having said their blood should be used for studies only, having more partners than other participants in the month preceding the initial interview, and participating in unsafe sex before the initial interview.

Discussion

To our knowledge, this is the first study that has evaluated behavior changes in a heterogeneous group of persons who

had not volunteered for HIV screening. Although every donor should have been notified that their blood would be tested for HIV antibodies, very few donated blood for the explicit purpose of being tested and many were unaware that the testing was done.[13] The results reported here provide insights into the types of reactions that other screened populations, such as immigrants[14] and hospitalized patients,[2] may experience.

A surprisingly high number of participants did not report being subject to a major HIV risk factor. We do not think this implies an unrecognized mode of transmission. Rather, we think many individuals do not recognize or acknowledge their risky behaviors. Many public health messages may not be optimally effective because they emphasize high-risk groups instead of high-risk behaviors. Our finding that many men who described themselves as exclusively heterosexual also reported having had sex with another man emphasizes that many individuals who engage in risky behaviors do not identify themselves with the associated high-risk groups. It may also be true that many individuals deny their personal risk even though they are aware of behaviors that increase the risk of HIV infection.[15] Much more needs to be known about the factors related to self-awareness of HIV risk.

The reductions in risky behaviors reported here are greater than those reported by some other investigators.[6,8;cf. 3,7] This may be due to the counseling that was provided, the self-selecting nature of those who enrolled in the study, or both.[8] It may also be true that sexual activity decreases after seroconversion because infection is most likely to occur during a peak of sexual activity.[16] We had no way of verifying reports of behavior, and participants may have underreported their risky behaviors at the follow-up interview.

Even if the reported reductions in risky behaviors were true and were due to the counseling provided, it is disconcerting that about 40% of the participants still reported having engaged in unsafe activities at the follow-up interview. This persistence in unsafe behavior by individuals who were given personal counseling and were motivated to join a psychosocial research project emphasizes the difficulty of encouraging people to adopt safer sexual behaviors. Whether long-term programs would be more successful at changing behavior remains to be determined.

Of the many psychosocial characteristics studied, only previous sexual behavior and use of the confidential self-exclusion option predicted the continuance of unsafe sexual behavior at the time of the follow-up interview; persons who engaged in high-risk behaviors and were aware of their risk were less likely to change their behavior. This finding is similar to that of McCusker and colleagues,[6] who found that previous sexual practices were the strongest predictors of subsequent practices. A major focus of our study was to learn more about the psychological and social characteristics that are related to beneficial behavior changes and to identify individuals who would benefit most from psychosocial support programs. Our lack of findings in these areas is discouraging. We examined only short-term behavior changes, and the effects of

TABLE 4—Sexual Behavior Reported at Initial Interview

	Men (%)	Women (%)
Number of sex partners of same sex in past year		
None	36.9	78.7
One	16.0	8.5
Two to five	27.8	8.5
More than five	19.3	4.3
Number of sex partners of opposite sex in past year		
None	58.6	33.3
One	18.3	33.3
Two to five	17.2	31.3
More than five	5.9	2.1
Number of sex partners of same sex in past month		
None	60.0	91.5
One	25.8	6.4
Two to five	13.7	2.1
More than five	0.5	0.0
Number of sex partners of opposite sex in past month		
None	77.0	39.6
One	16.6	47.9
Two to five	5.3	12.5
More than five	1.1	0.0
Percent of sexually active participants who always use condoms	39.6	42.2

TABLE 5—Sexual Behavior Reported at Initial Interview and at 2-Week Follow-Up

	Initial Interview (n = 271)		2-Week Follow-Up (n = 196)	
	Men (%)	Women (%)	Men (%)	Women (%)
Sexually active in previous week	88.5	86.5	83.6*	68.9*
Engaged in unsafe sex in previous week	68.1	57.7	40.4*	37.8**

Note. Significance tests were based only on information from persons who completed both questionnaires (n = 196).
*$P < .001$.
**$P < .01$.

TABLE 6—Logistic Regression Model of Unsafe Sex Reported at Follow-up (n = 196)

	Unsafe Sex Reported at Follow-up
Donated blood for studies only	.21*
Number of partners in month preceding initial interview	.25*
Unsafe sex reported at initial interview	.43***

Note. Coefficients are standardized logistic regression coefficients.
*$P < .05$; **$P < .01$; ***$P < .001$.

these characteristics may be less evident on short-term than on long-term behavior changes. There have been several excellent studies of the psychological and social factors related to behaviors that increase the risk of HIV infection,[6,8,17–20] but we still need to know much more about such factors. We currently are analyzing the relationship between various individual characteristics and behavior changes over longer periods of time.

Only about 38% of detected seropositive donors enrolled in the study, and only 28% completed both waves of data collection. We think it is likely that those who participated in this study were the most concerned about their health and were the most prone to change their behavior. Such self-selection, which will be almost inevitable in any HIV screening and counseling program, makes it extremely difficult to promote safe behavior after HIV screening.

Several researchers have argued that we should exercise extreme caution when developing routine screening programs.[2,6,8,14,21] The data reported here emphasize the difficulty of promoting voluntary behavior changes in seropositive individuals. The short-term persistence of unsafe behavior in this sample emphasizes the need for long-term follow-up.[22] We are currently evaluating a randomized controlled trial of such a program with blood donors.[1,23] Unless we learn more about ways to maximize the effectiveness of screening programs in promoting the desired behavior changes, they are likely to be a costly but ineffective and inefficient method of slowing the spread of HIV infection. □

Acknowledgments

This research was supported by grants from the National Heart, Lung, and Blood Institute (HL 35403-0151) and American Foundation for AIDS Research (AmFAR). The authors acknowledge and express their appreciation for the work of the other nurse clinicians who provided support and counseling to the donors and made the research reported in this paper possible. They include Catherine Tendler and Nalda Brodegaard. The authors thank Mary Hendricks, William McMullen, and Cathy Simon for help with data management and analyses and Carol Gerrior for help with manuscript preparation. The authors and staff of the project express their greatest respect for and appreciation to the seropositive individuals who contributed their time and thoughts to this project.

References

1. Cleary PD, Rogers TF, Singer E, et al. Health education about AIDS among seropositive blood donors. *Health Educ Q.* 1986;13:317–329.
2. Brandt AM, Cleary PD, Gostin LO. Routine hospital testing for HIV: health policy considerations. In: Gostin LO, ed. *AIDS and the Health Care System.* New Haven, Conn: Yale University Press; 1990.
3. Coates TJ, Stall RD, Kegeles SM, Lo B, Morin SF, McKusick L. AIDS Antibody Testing: Will it stop the AIDS epidemic? Will it help people infected with HIV? *Am Psychol.* 1988;43:859–864.
4. Cates W Jr, Handsfield HH. HIV counseling and testing: does it work? *Am J Public Health.* 1988;78:1533–1534.
5. Fox R, Odaka NJ, Brookmeyer R, Polk BF. Effect of HIV antibody disclosure on subsequent sexual activity in homosexual men. *AIDS.* 1987;1:241–246.
6. McCusker J, Stoddard AM, Mayer KH, Zapka J, Morrison C, Saltzman SP. Effects of HIV antibody test knowledge on subsequent sexual behaviors in a cohort of homosexually active men. *Am J Public Health.* 1988;78:462–467.
7. Wiktor SZ, Biggar RJ, Melbye M, et al. Effect of knowledge of human immunodeficiency virus infection status on sexual activity among homosexual men. *J Acquired Immune Deficiency Syndromes.* 1990;3:62–68.
8. Ostrow DG, Joseph JG, Kessler R, et al. Disclosure of HIV antibody status: behavioral and mental health correlates. *AIDS Educ Prev.* 1989;1:1–11.
9. van Griensven GJP, de Vroome EMM, Tielman RAP, et al. Effect of human immunodeficiency virus (HIV) antibody knowledge on high-risk sexual behavior with steady and nonsteady sexual partners among homosexual men. *Am J Epidemiol.* 1989;129:596–603.
10. Pindyck J, Waldman A, Zang E, Oleszko W, Lowy M, Bianco C. Measures to decrease the risk of acquired immunodeficiency syndrome transmission by blood transfusion: evidence of blood donor cooperation. *Transfusion.* 1985;25:3–9.
11. Cleary PD, Van Devanter N, Rogers TF, et al. Coping strategies and depressive symptoms in blood donors notified of HIV infection. Unpublished manuscript.
12. Nunnally JC. *Psychometric Theory.* New York: McGraw-Hill; 1978.
13. Cleary PD, Van Devanter N, Rogers TF, Singer E, Avorn J, Pindyck J. Trends in sociodemographic and behavioral characteristics of HIV antibody-positive blood donors. *AIDS Educ Prev.* 1991;3:60–71.
14. Gostin LO, Cleary PD, Mayer KH, Brandt AM, Chittenden EH. Screening immigrants and international travelers for the human immunodeficiency virus. *N Engl J Med.* 1990;322:1743–1746.
15. Cleary PD. Why people take precautions against health risks. In: Weinstein ND, ed. *Taking Care: Understanding and Encouraging Self-protective Behavior.* New York: Cambridge University Press, 1987.
16. Kuiken CL, van Griensven GJP, de Vroome EMM, Coutinho RA. Risk factors and changes in sexual behavior in male homosexuals who seroconverted for human immunodeficiency virus antibodies. *Am J Epidemiol.* 1990;132:523–530.
17. Emmons CA, Joseph JG, Kessler RC, Wortman CB, Montgomery SB, Ostrow DG. Psychosocial predictors of reported behavior change in homosexual men at risk for AIDS. *Health Educ Q.* 1986;13:331–345.
18. Stall RD, Coates TJ, Hoff C. Behavioral risk reduction for HIV infection among gay and bisexual men: a review of results from the United States. *Am Psychol.* 1988;43:878–885.
19. Becker MH, Joseph JG. AIDS and behavioral change to reduce risk: a review. *Am J Public Health.* 1988;78:394–410.
20. Martin JL, Dean L, Garcia M, Hall W. The impact of AIDS on a gay community: changes in sexual behavior, substance use, and mental health. *Am J Community Psychol.* 1989;17:269–293.
21. Cleary PD, Barry MJ, Mayer KH, Brandt AM, Gostin L, Fineberg HV. Compulsory premarital screening for the human immunodeficiency virus. *JAMA.* 1987;258:1757–1762.
22. Francis DP, Anderson RE, Gorman ME, et al. Targeting AIDS prevention and treatment toward HIV-1-infected persons: the concept of early intervention. *JAMA.* 1989;262:2572–2576.
23. Shipton-Levy R, Tendler C, Van Devanter N, Cleary PD. A group intervention model for individuals testing positive for HIV antibody. *Am J Orthopsychiatry.* 1990;60:452–459.

Knowledge about AIDS and HIV in the US Adult Population: Influence of the Local Incidence of AIDS

Linda F. McCaig, MPH, Ann M. Hardy, DrPH, and Deborah M. Winn, PhD

ABSTRACT

Background. Accurate information about acquired immunodeficiency syndrome (AIDS) and human immunodeficiency virus (HIV) is important for prevention. This study determined whether AIDS knowledge varied among populations residing in areas with a low, medium, or high incidence of AIDS.

Methods. Respondents to the 1988 National Health Interview Survey of AIDS Knowledge and Attitudes, which is based on a nationally representative sample of 29 659 adults, were rated on their knowledge about modes of HIV transmission, general knowledge about AIDS, and misperceptions about HIV transmission through casual contact.

Results. Persons 50 years of age or older, Blacks, Hispanics, and persons with less than a high school education had lower knowledge scores and higher misperception scores. Residents of the high-incidence area had more misperceptions than those who lived in the medium- or low-incidence areas. Sociodemographic determinants of scores were important factors overall and within each AIDS incidence area.

Conclusions. These data suggest that the ability of educational messages to reach and be absorbed by individuals nationwide is less dependent on whether they live in an area with a high incidence of AIDS and more dependent on other demographic variables. New and continued efforts are needed to improve knowledge in older persons, minorities, and the less educated in all parts of the country. (*Am J Public Health.* 1991;81:1591–1595)

Introduction

On the national and local levels, much effort in the control and prevention of human immunodeficiency virus (HIV) infection and acquired immunodeficiency syndrome (AIDS) has focused on educating both the general public and specific high-risk groups. It is important to evaluate these efforts periodically to determine if educational messages have reached their targets and to examine factors that influence knowledge levels. Evaluation of the national AIDS public information campaign has involved a variety of methods and data sources, including the National Health Interview Survey (NHIS) current health topic on AIDS Knowledge and Attitudes.

AIDS incidence rates in a geographic area may affect the public's knowledge about AIDS and HIV infection for several reasons. For one thing, demographic variables (e.g., education) that are associated with different levels of knowledge may be distributed differently among low-, medium-, and high-incidence areas. For another, local educational efforts may have begun earlier or been more intense in areas of higher incidence. Moreover, persons in high-incidence areas may have a greater chance of knowing someone with AIDS and/or may perceive themselves to be at greater personal risk of developing AIDS, either of which factor may increase their opportunity and desire to learn more about the disease. Finally, national polls have found that, as the distance from the major AIDS epicenters increases, the public perceives AIDS to be less threatening.[1] Thus, if knowledge of AIDS and HIV infection is affected by the incidence of AIDS in the area under study, this information should be incorporated into the planning of public information efforts.

This study examined data from the 1988 NHIS AIDS Knowledge and Attitudes survey to determine how knowledge varied by demographic characteristics and among populations residing in areas with low, medium, and high annual incidence rates of AIDS. Because the NHIS is the only survey of AIDS knowledge that is based on a national probability sample, it provided a unique opportunity to evaluate knowledge of AIDS and HIV infection by incidence areas on a national level.

Methods

The NHIS, a cross-sectional sample survey of the US civilian, noninstitutionalized population, has been conducted annually since 1957 by the National Center for Health Statistics, one of the Centers for Disease Control (CDC).[2] Each year, members of 50 000 households across the country are interviewed in person by US Bureau of the Census personnel. The NHIS sample design is a multistage, stratified, probability sample of the US population.[3]

The first questionnaire about AIDS was included in the NHIS in 1987. The questionnaire was then revised, and data collection with the version of the AIDS survey described in this report began in 1988. The questionnaire, administered to one randomly selected adult aged 18 years

At the time of the study the authors were with the Division of Health Interview Statistics, National Center for Health Statistics, in Hyattsville, Md.

Requests for reprints should be sent to Linda F. McCaig, MPH, Health Statistician, Ambulatory Care Statistics Branch, Division of Health Care Statistics, National Center for Health Statistics, 6525 Belcrest Road, Room 950, Hyattsville, MD 20782.

This paper was submitted to the journal October 10, 1990, and accepted with revisions May 6, 1991.

TABLE 1—Questions and Answers on Which the Transmission Knowledge, General Knowledge, and Misperception Scores Were Based

Question	Correct Answer[a]
Transmission Knowledge Scores	
ANY person with the AIDS virus can pass it on to someone else through sexual intercourse.	DT or PT
A pregnant woman who has the AIDS virus can give the AIDS virus to her baby.	DT or PT
How likely do you think it is that a person will get AIDS or the AIDS virus infection from sharing needles for drug use with someone who had the AIDS virus?	Very likely or somewhat likely
General Knowledge Scores	
AIDS is especially common in older people.	DF
AIDS usually leads to heart disease.	DF
AIDS leads to death.	DT
A person can be infected with the AIDS virus and not have the disease AIDS.	DT
ANY person with the AIDS virus can pass it on to someone else through sexual intercourse.	DT or PT
There is a vaccine available to the public that protects a person from getting the AIDS virus.	DF
There is no cure for AIDS at present.	DT
Misperception Scores[b]	
How likely do you think it is that a person will get AIDS or the AIDS virus from—	
living near a home or hospital for AIDS patients?	
working near someone with the AIDS virus?	
sharing plates, forks, or glasses with someone who has the AIDS virus?	
using public toilets?	
being coughed on or sneezed on by someone who has the AIDS virus?	
attending school with a child who has the AIDS virus?	
mosquitoes or other insects?	

[a]DT or PT = Definitely true or probably true; DF = Definitely false.
[b]The correct answer to all of the misperception questions is "Definitely not possible or very unlikely."

or older in each family included in the NHIS sample, was designed primarily to estimate adults' knowledge about AIDS, HIV transmission, and prevention of HIV infection. In the questionnaire, the term "AIDS virus" was used instead of "HIV" because it was believed to be more readily understood. In this report, however, "AIDS" refers to the disease, "HIV" to the virus.

Sample persons were classified by the annual incidence rate of AIDS in their county of residence. The annual incidence rates of AIDS in census-defined metropolitan statistical areas (MSAs) with 500 000 or more population for the period April 1988 to March 1989 was ascertained from the *HIV/AIDS Surveillance Report* published by CDC.[4] The 94 MSAs listed in the CDC report were then compared with the counties included in the NHIS sample.[5] The 6 MSAs not part of the NHIS sample were excluded, and the remaining 88 MSAs were ranked according to incidence rate from lowest to highest. Three groups were then determined by two break points in the list where the rates in one area differed from those in the next higher area by 1.5% or more. Also considered in making these determinations was that each incidence area would have a sufficient sample size for analyses. The incidence areas were classified as low (fewer than 10 cases per 100 000 population, 46 MSAs, range 1.9–9.6), medium (10 to 27 cases per 100 000 population, 35 MSAs, range 11.1–26.6), or high (more than 27 cases per 100 000 population, 7 MSAs, range 34.7–99.7).

Incidence rather than prevalence rates were used to classify counties because that was the measure available as reported by CDC. However, in 1988, ranking cities by prevalence or incidence rates would probably have yielded the same order. CDC provides annual AIDS incidence data only for MSAs with 500 000 or more population to ensure the anonymity of persons with AIDS who reside in small cities or rural areas.[4] Forty-one percent of the NHIS sample did not live in the MSAs listed in the CDC report; therefore, it was not possible to classify all counties in the NHIS sample as to their incidence of AIDS.

In 1988, 29 659 adults—or about 90% of eligible sample persons—responded to the AIDS questionnaire; of those, 17 277 persons could be classified into an AIDS incidence area. This includes 58% of the NHIS sample eligible for the AIDS Knowledge and Attitudes survey and represented 59% of the US population aged 18 years and over. For results presented by AIDS incidence area, only the part of the sample with incidence assigned was used; for national estimates, the entire NHIS sample eligible for the AIDS Knowledge and Attitudes survey was used.

Three scales reflecting AIDS knowledge were constructed: transmission knowledge, general knowledge, and misperception scores. The 1988 NHIS AIDS questionnaire contained three questions that were used in this report to assess knowledge about recognized modes of HIV transmission (i.e., sexual, perinatal, and parenteral) (Table 1). For each respondent, the number of correct responses to these three statements was summed to yield a transmission knowledge score; this score could range from 0 to 3, with higher scores indicating more knowledge. Seven statements were used to assess general AIDS knowledge. For each respondent, the number of correct responses to these seven statements was summed to produce a general knowledge score; this score could range from 0 to 7, with higher scores indicating more knowledge. Seven questions were used to assess misperceptions about HIV transmission through casual contact. For each respondent, the number of incorrect responses was summed to produce a misperception score; this score could range from 0 to 7, with higher scores indicating more misperceptions about HIV transmission through casual contact.

The demographic variables examined included age (18 to 49 years and 50 years or older), race ethnicity (White non-Hispanic, Black non-Hispanic, and Hispanic), and education, given that knowledge differences within these categories have been previously reported from the NHIS current health topic on AIDS Knowledge and Attitudes[6,7] and elsewhere.[1,8–14] Data were also analyzed by marital status for males younger than 50 years old because most AIDS cases have occurred in young men.[15] Single men

included those who were divorced, separated, or never married.

To make national estimates, estimates derived from the NHIS sample were weighted to take into account selection probability, nonresponse, and an adjustment to the current US population.[3] Mean scores and proportions were generated with SAS software. Standard errors (SEs) and statistical differences between mean scores were calculated with SESUDAAN, which takes into account the complex sample design of the NHIS.[16] For the comparison of two means, the standard error of the difference (SED) was used to calculate the 95% confidence interval around the difference. All the differences noted in the text were significant at the 0.05 level.

Results

Of the respondents who could be classified into an AIDS incidence area, 12% (n = 2049) resided in an area with a high annual incidence of AIDS, 50% (n = 8625) resided in a medium-incidence area, and 38% (n = 6603) resided in a low-incidence area. Thirty-seven percent of the persons who lived in a high-incidence area were 50 years of age or older compared with 32% and 34% in medium- and low-incidence areas, respectively. Blacks made up 25% of the study group who lived in a high-incidence area compared with 16% in both medium- and low-incidence areas. Eighteen percent of persons residing in a high-incidence area were Hispanic compared with 10% and 4% in medium- and low-incidence areas, respectively. Twenty-three percent of persons living in a high-incidence area had less than a high school education compared with 19% and 21% in medium- and low-incidence areas, respectively. Fifty-one percent of unmarried males who were younger than 50 years of age resided in a high-incidence area compared with 47% and 44% in medium- and low-incidence areas, respectively.

When the estimated overall mean transmission knowledge, general knowledge, and estimated misperception scores for persons classified into the three incidence groupings were compared with estimated scores for persons in areas not classified, no differences were found.

Transmission Knowledge Scores

The estimated mean transmission knowledge score was 2.8 (SE = .01) out of a possible 3 for both the nation and each AIDS incidence grouping. Persons 50 years of age or older had a mean transmission knowledge score of 2.7 compared with 2.9 for those aged 18 to 49 years (SED = .01). Those with less than a high school education had a lower mean score (2.6) than those who had 12 years of education (2.9, SED = .01) and those who had 12 or more years of school (2.9, SED = .01). No differences in scores were noted by AIDS incidence area. SEDs for mean transmission knowledge scores in subgroups within incidence groupings ranged from .03 to .06.

TABLE 2—Mean General Knowledge Scores by Demographic Characteristics and AIDS Incidence Groupings

Demographic Characteristics	General Knowledge Scores[a] AIDS Incidence Groupings: High	Medium	Low	US Total	Maximum Differences in Scores
Total	4.7	4.8	4.8	4.8	0.1
Age:					
18–49 years	4.9	5.0	5.0	5.0	0.1
50+ years	4.4	4.4	4.4	4.4	0.0
Race:					
White	4.7	4.8	4.8	4.7	0.1
Black	4.4	4.4	4.5	4.3	0.1
Hispanic	4.6	4.3	4.3	4.4	0.3
Education:					
<12 years	4.0	4.1	4.1	4.0	0.1
12 years	4.7	4.8	4.9	4.8	0.2
>12 years	5.2	5.1	5.1	5.1	0.1
Marital status:[b]					
Single males	4.6	4.8	4.7	4.7	0.2
Married males	4.9	4.8	5.0	4.9	0.2

[a]Mean number of correct responses out of a possible seven.
[b]Includes males younger than 50 years old.

General Knowledge Scores

The estimated mean general knowledge score for the nation was 4.8 (SE = .02) out of a possible 7, and overall, scores were nearly identical in each incidence grouping (Table 2). However, for the nation and within each incidence grouping, differences were observed for age, race ethnicity, education, and marital status for males. By level of education, the racial/ethnic differences remained except for Hispanics with 12 or more years of school. Hispanics who lived in a high-incidence area had a higher mean general knowledge score (4.6) than those in areas of medium (4.3, SED = .10) or low (4.3, SED = .14) incidence. SEDs for mean general knowledge scores in subgroups within incidence groupings ranged from .04 to .14.

Misperception Scores

The estimated mean misperception score for the nation was 2.9 (SE = .03) out of a possible 7. Overall and within each incidence grouping, differences were observed for age, race ethnicity, education, and marital status for males (Table 3). The effect of race/ethnicity on misperception scores remained after controlling for education except for Hispanics with 12 years of school.

Overall, persons who lived in a high-incidence area had a higher mean misperception score (3.1) than those who lived in areas of medium (2.8, SED = .09) or low (2.8, SED = .09) incidence. Persons 50 years of age or older who lived in a high-incidence area had a higher mean misperception score (3.6) than those in the same age group who lived in areas of medium (3.3, SED = .13) or low (3.3, SED = .13) incidence. This pattern was not observed in younger adults. Persons who had 12 years of education and who lived in a high-incidence area had a higher average score (3.3) than persons with the same number of years of education who lived in areas of medium (3.0, SED = .12) or low (2.9, SED = .08) incidence, but no trends were observed among those with more or with less education. Single men in a low-incidence area had a higher average misperception score (2.8) than single men in a medium-incidence area (2.5, SED = .14). Married men in a high-incidence area had a higher average misperception score (3.0) than married men in a low-incidence area (2.6, SED = .16). SEs for mean misperception scores in subgroups within incidence groupings ranged from .06 to .25.

TABLE 3—Mean Misperception Scores by Demographic Characteristics and AIDS Incidence Groupings

Demographic Characteristics	Misperception Scores[a] AIDS Incidence Groupings				
	High	Medium	Low	US Total	Maximum Difference in Scores
Total	3.1	2.8	2.8	2.9	0.3
Age:					
18–49 years	2.8	2.6	2.6	2.7	0.2
50+ years	3.6	3.3	3.3	3.4	0.3
Race:					
White	2.8	2.6	2.7	2.8	0.2
Black	3.4	3.4	3.4	3.5	0.0
Hispanic	3.2	3.4	3.4	3.3	0.2
Education:					
<12 years	4.1	3.9	3.9	4.0	0.2
12 years	3.3	3.0	2.9	3.0	0.4
>12 years	2.3	2.3	2.2	2.3	0.1
Marital status:[b]					
Single males	2.6	2.5	2.8	2.7	0.3
Married males	3.0	2.8	2.6	2.8	0.4

[a]Mean number of incorrect responses out of a possible seven.
[b]Includes males younger than 50 years old.

To test whether scores for counties with a very high annual incidence rate of AIDS were in any way remarkable, the three MSAs with the highest incidence rates were compared with the rest of the classifiable counties. The mean transmission knowledge scores were identical (2.8). However, the mean general knowledge score for the three highest MSAs (4.5) was significantly lower than that for the remainder of the counties (4.7, SE = .04), while the mean misperception score for the three highest MSAs (3.3) was significantly higher than that for the other counties (2.8, SE = .11).

Discussion

An effective educational program is a crucial part of the public health response to AIDS. The information from the continuing NHIS AIDS Knowledge and Attitudes survey has been useful for guiding national educational efforts and for evaluating specific campaigns such as the CDC's "Understanding AIDS" brochure, which was mailed to every US household in spring 1988.[17]

Disease knowledge and risk perceptions in an area might be hypothesized to be either positively or negatively associated with disease incidence in a cross-sectional survey such as the NHIS. Ignorance of disease risk factors and of preventive measures might lead to adoption or continuance of behaviors that place an individual at risk, which, in turn, results in high incidence rates for the disease. Alternatively, in areas where a disease is common, effective public health efforts to educate the population could lead to an increasing knowledge of the disease. Misperceptions about transmission might be exacerbated by the disease frequency in an area if they are fueled by the public's concern about the greater probability of coming in contact with infected persons in a high-incidence area. On the other hand, misperceptions may be lower in a high-incidence area if public health education has been successful.[18]

This analysis of the 1988 NHIS AIDS Knowledge and Attitudes survey shows that knowledge about the major modes of HIV transmission (i.e., sexual, parenteral, and perinatal) was high in all demographic groups in the United States. Overall, the nation's general knowledge about AIDS and HIV infection was high, and misperceptions about the risk of transmitting the virus through casual contact were few. This confirms data from numerous national, state, and local studies.[1,6,7,12,14,19–22]

Previous studies have identified older persons, Blacks, Hispanics, and those with fewer than 12 years of education as being less knowledgeable about AIDS,[1,6,8–14] a finding also supported by our study. Our report further demonstrates that the demographic determinants (i.e., age, race ethnicity, education) of knowledge and the misperception scores were the same within each AIDS incidence area as they were nationally.

Among persons living in a high incidence area for AIDS, higher misperception scores were noted overall and for several subgroups: persons 50 years of age or older, persons with 12 years of education, and married men. Some of these differences may be explained by the proportionately higher number of older persons, minorities, and persons with less than a high school education who resided in high-incidence areas. However, the absolute differences in misperception scores among those living in areas of high, medium, or low AIDS incidence were modest compared with differences attributable to age, race ethnicity, and education. In addition, because of the large sample size of the NHIS, some differences that were statistically significant were in magnitude quite small and probably of little practical importance. We did note larger differences when we focused on areas of extremely high AIDS incidence; however, these differences may have been due more to the distribution of certain demographic characteristics than to incidence. An inadequate sample size prevented us from further stratifying the data. National public opinion surveys conducted between 1983 and 1986 also found that misperceptions about AIDS showed some geographical variation.[1] Perhaps these differences lessened by 1988, when the NHIS data used in this report were collected.

Our overall findings of no differences in transmission and general knowledge scores in areas that differed markedly in the annual incidence of AIDS is in contrast to a recent finding in a study of gay men, in which bar clientele in areas with high AIDS prevalence were more knowledgeable about AIDS than men in low-prevalence areas.[18] Homosexual men are at especially high risk, and findings for them may not apply to other groups. The NHIS AIDS Knowledge and Attitudes survey does include a general question assessing whether respondents have engaged in any of various HIV risk behaviors; however, only 2% reported engaging in such behaviors, and this group was too small to analyze in this study.

Because AIDS incidence data are not available for all US counties, we could not classify 41% of the NHIS sample. Therefore, the findings from this report may not apply directly to these areas. However, areas not included are disproportionately nonmetropolitan (57%) and presumably of low incidence. Approximately 25% of nonclassifiable areas are in states with low incidence rates (fewer than 5 cases per 100 000 population), even in metropolitan

areas. In addition, general knowledge about AIDS in areas that were not classifiable did not differ from such knowledge in areas that could be classified; the latter included areas with a wide range of population densities and presumably a wide cross-section of the American public.

The observation that AIDS transmission knowledge and general knowledge scores in this national survey are similar for various demographic groups in the United States, regardless of the incidence level of AIDS, suggests that the ability of educational messages to reach and be absorbed by individuals nationwide depends less on whether those individuals currently live in an area where the annual incidence of AIDS is high than it does on demographic and other variables. Our finding, which shows the influence of age, race, ethnicity, and education on knowledge and perceptions about AIDS in geographic areas throughout the United States that vary by threefold or more in annual AIDS incidence, reinforces the importance of new and continued efforts to improve knowledge among older persons, minorities, and the less educated in all parts of the country. □

Acknowledgments

The authors would like to thank Steven L. Botman, MA, and Ronald W. Wilson, MA, for thoughtfully reviewing the manuscript.

References

1. Blake SM, Arkin EB. A summary of national public opinion surveys on AIDS: 1983 through 1986. AIDS Information Monitor. Washington, DC: American Red Cross, 1986.
2. Kovar MG, Poe GS. *The National Health Interview Survey Design, 1973–84, and Procedures, 1975–1983.* Hyattsville, Md: National Center for Health Statistics, Vital and Health Statistics; 1985:1(18).
3. Massey JT. *Design and Estimation for the National Health Interview Survey, 1985–94.* Hyattsville, Md: National Center for Health Statistics, Vital and Health Statistics; 1989:2(110).
4. Centers for Disease Control. *HIV/AIDS Surveillance Report:* April 1989, 6–7.
5. US Department of Commerce, Bureau of the Census. PSU's included in current sample surveys. USGPO, Washington, DC: September 1986.
6. Hardy AM, Dawson DA. *AIDS Knowledge and Attitudes for December 1988: Provisional Data from the National Health Interview Survey.* Advance data from vital and health statistics; no. 175. Hyattsville, Md: National Center for Health Statistics; 1989.
7. Centers for Disease Control. HIV epidemic and AIDS: trends in knowledge—United States, 1987 and 1988. *MMWR.* 1989;38: 353–363.
8. DiClemente RJ, Boyer CB, Morales ES. Minorities and AIDS: knowledge, attitudes, and misconceptions among Black and Latino adolescents. *Am J Public Health.* 1988;78:55–57.
9. St. Lawrence JS, Betts R. Comparison of the knowledge, attitudes, and AIDS-risk behavior of Black and White college students in the Southeastern United States. Presented at the Fifth International Conference on AIDS. Abstracts; 1989:747.
10. Seltzer R, Smith RC. Racial differences and intraracial differences among Blacks in attitudes towards AIDS. *AIDS Public Policy J.* 1988;3:31–35.
11. Sonenstein FL, Pleck JH, Ku LC. Sexual activity, condom use and AIDS awareness among adolescent males. *Fam Plann Perspect.* 1989;21:152–158.
12. Sy FS, Freeze-McElwee Y, Garrison CZ, Jackson KL. Knowledge, perceived risk, and beliefs about AIDS among high school and college students in South Carolina. *J SC Med Assoc.* 1989;481–486.
13. Kappel S, Vogt RL, Brozicevic M, Kutzko D. AIDS knowledge and attitudes among adults in Vermont. *Public Health Rep.* 1989;104:388–391.
14. Keeter S, Bradford JB. Knowledge of AIDS and related behavior change among unmarried adults in a low-prevalence city. *Am J Prev Med.* 1988;4:146–152.
15. Centers for Disease Control. AIDS and human immunodeficiency virus infection in the United States: 1988 update. *MMWR.* 1988;38(Suppl. S–4):1–38.
16. Shah BV. *SESUDAAN: Standard Error Program for Computing Standardized Rates from Sample Survey Data.* Research Triangle Park, NC: Research Triangle Institute; 1981.
17. Rosenberg M, Dawson DA, Cyndmon ML. An evaluation of the national mailout "Understanding AIDS." Presented at the annual meeting of the American Public Health Association; November 13–17, 1988; Boston, Mass.
18. St. Lawrence JS, Hood HV, Brasfield T, Kelly JA. Differences in gay men's AIDS risk knowledge and behavior patterns in high and low AIDS prevalence cities. *Public Health Rep.* 1989;104:391–395.
19. McKusick L, Horstman W, Coates TJ. AIDS and sexual behavior reported by gay men in San Francisco. *Am J Public Health.* 1985;75:493–496.
20. Kinnick BC, Smart DW, Bell DA, Blank WR, Gray TR, Schober JL. An assessment of AIDS-related knowledge, attitudes, and behaviors among selected college and university students. *AIDS Public Policy J.* 1989;4:112–119.
21. Goodman E, Cohall AT. Acquired immuno-deficiency syndrome and adolescents: knowledge, attitudes, beliefs, and behaviors in a New York City adolescent minority population. *Pediatrics.* 1989;84:36–42.
22. Beaman ML, Strader MK. STD patients' knowledge about AIDS and attitudes toward condom use. *J Community Health Nurs.* 1989;6:155–164.

HIV Instruction, HIV Knowledge, and Drug Injection among High School Students in the United States

Deborah Holtzman, PhD, John E. Anderson, PhD, Laura Kann, PhD, Susan L. Arday, MHS, Benedict I. Truman, MD, MPH, and Lloyd J. Kolbe, PhD

ABSTRACT

Background. The prevalence of HIV-related behaviors and determinants of these behaviors among adolescents in the United States have not been well studied.

Methods. To determine the prevalence of HIV-related drug behaviors and to assess the effects of HIV-related school-based instruction and HIV knowledge on these behaviors, data were analyzed from a 39-item, self-administered questionnaire completed by a probability sample of all students in grades 9 through 12 in the United States.

Results. Usable responses were obtained from 8098 students. Of these, 2.7% (95% confidence interval [CI] = 2.3–3.2) and 1.7% (95% CI = 1.3–2.1) reported injecting illicit drugs ever and during the past year, respectively. Corresponding prevalences of needle sharing were 0.8% (95% CI = 0.5–1.1) and 0.5% (95% CI = 0.3–0.7). Regression analysis revealed that students with higher knowledge scores were less likely and males more likely to have ever injected drugs. HIV knowledge was similarly associated with other outcome measures of drug-injection behavior. Although HIV instruction did not directly influence drug-injection behavior independently of demographic characteristics, it was positively associated with HIV knowledge.

Conclusions. While these results do not establish a causal relationship, they suggest that HIV knowledge and school-based instruction may play a role in maintaining low levels of drug-injection behavior among high school students. (*Am J Public Health.* 1991;81:1596–1601)

Introduction

This article describes the prevalence of self-reported illicit drug injection among high school students in the United States and examines the relationships among school-based instruction about human immunodeficiency virus (HIV), HIV knowledge, and drug-injection behavior. Although the incidence of acquired immunodeficiency syndrome (AIDS) in US adolescents is not as high as it is in adults and children, each of these groups has experienced a rapid rate of growth in the cumulative incidence of the disease in the last decade.[1,2] Because of the long latency period for HIV, which may average 10 years or more,[3] it is also likely that many of the 20- to 29-year-olds diagnosed with AIDS, who account for 20% of all cases nationwide,[4] were infected with HIV during adolescence. Indeed, many adolescents engage in behaviors that place them at risk for HIV infection. As of July 1990, 40% of the AIDS cases in the United States reported among 13- to 19-year-olds resulted from sexual transmission of HIV, 11% of cases reported histories of intravenous (IV) drug use, and another 4% reported both risk behaviors.[4]

In national surveys, reported rates of sexual intercourse among 17- to 19-year-old, never-married males in metropolitan areas increased from 65.7% in 1979 to 75.5% in 1988.[5] In similar surveys of US females, the rates of premarital sexual intercourse among 15- to 19-year-old adolescents living in metropolitan areas increased from 43.4% in 1979 to 45.2% in 1982.[6] Further, a 1979 survey of adolescent females revealed that their mean age at first intercourse was 16.2 years and that 16% of the sample reported four or more sexual partners.[7,8] In a 1989 nationwide survey, 58.5% of 9th- through 12th-grade students had had sexual intercourse, and of these, 40.4% reported four or more partners. In addition, 23.3% of the students who ever had sex reported that they never used a condom when they had sexual intercourse.[9]

IV drug use has also been reported in surveys of youth, but the extent and nature of illicit drug-injection behavior among adolescents has not been well studied. Surveys of IV drug use in relatively small or geographically limited populations of adolescents have shown considerable variation in rates, ranging from 0.1 to 6.3%.[10–15] How prevalent, then, is IV drug use among adolescents in the United States? And equally important, what factors influence this behavior?[16]

Several surveys of HIV-related knowledge, attitudes, and behaviors among adolescents have been reported,[5,10–15,17,18] but only one was nationwide in scope,[13]

Deborah Holtzman, Laura Kann, Susan L. Arday, Benedict I. Truman, and Lloyd J. Kolbe are with the Division of Adolescent and School Health, National Center for Chronic Disease Prevention and Health Promotion, and John E. Anderson is with the Division of STD/HIV Prevention, National Center for Prevention Services, all at the Centers for Disease Control, Atlanta, Ga.

Requests for reprints should be sent to Deborah Holtzman, PhD, Division of Adolescent and School Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control, Mailstop K-33, 1600 Clifton Road, NE, Atlanta, GA 30333.

This paper was submitted to the journal November 15, 1990, and accepted with revisions April 4, 1991.

Editor's Note. See related Editorial by Hinman on page 1557.

only a few assessed risk behaviors for HIV infection,[10-12,14,15] and fewer still have analyzed possible determinants of these behaviors.[11]

As part of a system of surveys for collecting school-based data, the Centers for Disease Control conducted the Secondary School Student Health Risk Survey (SSSHRS) in April and May of 1989. The SSSHRS was the first survey to collect data about HIV-related knowledge, beliefs, and behaviors among a nationally representative sample of US students in grades 9 through 12. Two variables that may be possible determinants of behavioral risk reduction—HIV instruction and HIV knowledge—were measured by the survey. The SSSHRS also collected data about sexual activity and illicit drug injection. This report focuses on the latter.

First, we examined the prevalence of injected drug use among high school students. Second, we examined whether students who had received classroom instruction about HIV and who had a greater level of HIV knowledge were less likely than other students to have injected drugs or shared needles. With this information we can better plan and evaluate our HIV educational efforts for adolescents.

Methods

Study Population

The sampling frame, which consisted of all US public- and private-school students in grades 9 through 12, was divided into 1921 primary sampling units (PSUs) and stratified by region and urbanicity. Fifty PSUs were selected and, from these PSUs, 50 counties were selected. From the 50 counties, 122 schools were selected. At each stage, the sample was selected with probability proportional to the size of the 9th through 12th grade enrollment. Heavily Black and Hispanic areas were oversampled. Within participating schools, one or two classes of a required subject were randomly selected.

Data Collection

Representatives from numerous federal, state, and local education agencies and experts in HIV and reproductive health behaviors were involved in developing the survey instrument. Personnel were trained to coordinate on-site data collection, and special measures were taken to protect student anonymity. A 39-item questionnaire was designed to be self-administered in the classroom. Five broad groups of variables were assessed by the questionnaire: demographic characteristics, HIV instruction, HIV attitudes, HIV knowledge, and HIV-related risk behaviors. Variables used in the current analysis are described in more detail below.

HIV Instruction, Knowledge, and Behavior. A single question, answered yes or no, was used to measure HIV instruction. Students were asked, "Have you been taught about AIDS/HIV infection in school?" Seventeen items measured HIV knowledge. Students were asked about such factors as modes of HIV transmission and how to reduce risk for HIV infection. Based on the number of correct responses to the 17 items, a score was computed for each respondent, with equal weight given to each question. Drug-injection behavior was assessed by four dichotomous dependent variables. Students were asked if they had done the following: (1) ever injected illicit drugs, (2) injected illicit drugs in the last year, (3) ever shared needles, or (4) shared needles in the last year.

Data Analysis

Two separate multivariate analyses were conducted to examine relationships among HIV knowledge, HIV instruction, and drug-injection behavior. Using multiple linear regression, HIV knowledge was regressed on HIV instruction, while controlling for age, sex, racial or ethnic group, and the interactions between HIV instruction and sex, and between HIV instruction and racial or ethnic group. Using stepwise multiple logistic regression, each of the four drug-behavior variables was regressed on HIV instruction and HIV knowledge, controlling for the same background characteristics. There are theoretical and empirical reasons for including both HIV instruction and knowledge in the models. Theoretically, for instruction to have an effect on drug behavior it would likely work through some other mechanism, such as knowledge or attitudes. Without knowledge, then, the model would be improperly specified, and would bias the effect of instruction on drug behavior. In addition, a comparison of the model with knowledge and without knowledge showed that the effect of HIV instruction on drug behavior did not change. A further complication arises, however, if HIV instruction and HIV knowledge are highly correlated. This was examined by the authors and the correlation between the two variables was found to be relatively low ($r = .21$), suggesting that multicollinearity was not a problem in the analysis.

Since very few students reported having shared needles within categories of racial or ethnic group and HIV instruction, the effect of these interactions on sharing needles could not be evaluated. Additionally, because the interactions between knowledge and each demographic characteristic and between HIV instruction and age were all highly correlated ($r > .82$) with at least one of the main effects, these terms were excluded from all models. All main effects, regardless of the level of significance, were retained in the models when the interaction term was significant.[19]

All data were weighted to compensate for nonresponse within categories of age, grade, and sex, and the fact that not all students were sampled with the same likelihood. The SAS programming

TABLE 1—Region, Area, and Mean Proportion of Racial or Ethnic Groups for Sampled Schools and Participating Schools, 1989 Secondary School Student Health Risk Survey

Category	Sampled (n = 122)	Participating (n = 99)
	No. of Schools (%)	
Region		
Northeast	17 (13.9)	14 (14.1)
North Central	27 (22.1)	20 (20.2)
South	53 (43.4)	47 (47.5)
West	25 (20.5)	18 (18.2)
Area		
Large metropolitan area	54 (44.3)	43 (43.4)
Other	68 (55.7)	56 (56.6)
	Mean Proportion	
Racial or ethnic group		
White	60.9	58.4
Black	11.6	13.2
Hispanic	11.0	11.6

language was used to conduct the analyses,[20] and statistical tests and confidence intervals were based on the procedures SURREGR, SESUDAAN, and RTILOGIT, which take into account the complex sample design.[21–23] Results were considered significant if the *P* value for the calculated statistic was less than or equal to .05.

Results

Of the 122 schools selected, 99 (81%) participated. Selected schools and participating schools were compared by region, metropolitan or nonmetropolitan area, and racial or ethnic group (Table 1). Only small differences were observed within each category, suggesting that school-level nonresponse did not greatly bias the sample.

Of selected students, 10% did not receive parental permission to participate and an additional 7% were absent when the survey was administered and on the scheduled make-up day(s). Usable questionnaires were received from 8098 (83%) of the eligible sample. Based on the school- and student-level response, an overall response rate of 67% was obtained.

Participants were fairly evenly distributed by grade and sex (Table 2). The weighted percentages of White, Black, and Hispanic students who participated were similar to estimates of the racial and ethnic distribution of 9th- through 12th-grade students in the United States. For the US population of students in these grades, the estimates were 70.8% White, 15.5% Black, and 8.9% Hispanic.[24] (Calculations for general population estimates of high school students were also based on unpublished data from US Bureau of the Census, Annual Projections of the Population by Age, Sex, Race, for the United States: 1983 to 2080, Hispanic by race data, July 1, 1988.)

Slightly over half (53.7%) of those surveyed said they had received instruction in school about AIDS and HIV infection (Table 3). Significant differences in the proportion of students reporting HIV instruction were found by age and grade only. Students who were 17 or more years of age were less likely than 15- and 16-year-old students to have received HIV instruction ($P < .05$), and 12th-grade students were less likely than 10th- or 11th-grade students to have received instruction ($P < .05$).

For all students, an average of 14.5 knowledge items were answered correctly (range = 0–17). The vast majority of students knew that AIDS/HIV infection is transmitted during sexual intercourse (97.7%), by sharing needles (98.9%), and perinatally (90.7%). Most also knew that HIV infection is not transmitted by casual contact. Students were less informed about transmission of the virus in other ways; only 64.0% knew the virus could not be transmitted by donating blood, and only 45.1% knew the virus could not be transmitted by insect bites.

When asked about drug-injection behavior, 2.7% of students reported having ever injected illicit drugs, and 0.8% reported having ever shared needles (Table 4). The rate of ever injecting drugs was higher for males than for females (3.7% vs 1.8%, $P < .05$). Differences in drug-injection behavior by racial or ethnic group were not statistically significant.

The reported rates of ever injecting drugs showed a clear decline from grades 9 to 12, and significant differences were observed between 9th- and 12th-grade students ($P < .05$). No significant differences, however, were found by age group. Finally, students who reported having been taught about HIV were less likely than those who reported no instruction to have ever injected drugs ($P < .05$).

Results of logistic regression analysis showed that HIV knowledge, being male, and the interactions between HIV instruction and being male and between HIV instruction and being Black all had significant effects on ever injecting drugs (Table 5). The greater the level of knowledge, the less likely students were to report ever injecting drugs. Males were more likely than females to have injected drugs, although males who had received HIV instruction were less likely to have injected

TABLE 2—Demographic Characteristics of Sample[a], 1989 Secondary School Student Health Risk Survey

Characteristic	Unweighted Frequency	Weighted Percentage
Age (years)		
<15	699	8.4
15–16	3958	48.9
>16	3438	42.7
Grade		
9	1916	23.9
10	2085	25.5
11	2049	25.5
12	2046	25.2
Racial or ethnic group		
White	5107	78.3
Black	1155	9.1
Hispanic	1368	7.9
Other[b]	461	4.7
Sex		
Female	4341	49.3
Male	3748	50.6

[a]N = 8098.
[b]Includes Native Americans, Alaskan Natives, Asian or Pacific Islanders, and persons of other unspecified groups.

TABLE 3—Percentage of Students Who Received HIV Instruction by Demographic Characteristic, 1989 Secondary School Student Health Risk Survey

Characteristic	Percentage (95%[a])
Age (years)	
<15	57.7 (49.6–65.8)
15–16	57.0 (50.9–63.1)
>16	49.1 (42.3–56.0)
Grade	
9	54.8 (47.6–61.9)
10	59.4 (50.6–68.1)
11	53.7 (46.6–60.8)
12	47.1 (39.4–54.7)
Racial or ethnic group	
White	54.0 (47.5–60.5)
Black	54.9 (47.2–62.7)
Hispanic	51.4 (43.1–59.6)
Other[b]	51.4 (41.5–61.2)
Sex	
Female	54.0 (48.0–60.1)
Male	53.4 (47.6–59.2)

Note. A total of 4270 (53.7%, 95% confidence interval = 48.0–59.5) received HIV instruction.
[a]CI = confidence interval.
[b]Includes Native Americans, Alaskan Natives, Asian or Pacific Islanders, and persons of other unspecified groups.

TABLE 4—Percentage of Students Who Reported Illicit Drug Injection and Needle Sharing by Demographic Characteristic and HIV Instruction, 1989 Secondary School Student Health Risk Survey

Characteristic	Ever Injected Drugs (n = 8009)[a]	Injected Drugs Last Year (n = 8009)[a]	Ever Shared Needles (n = 7954)[a]	Shared Needles Last Year (n = 8006)[a]
Total	2.7	1.7	0.8	0.5
Age (years)				
<15	2.2	1.6	0.9	0.5
15–16	3.1	2.1	0.7	0.3
>16	2.4	1.4	0.8	0.7
Grade				
9	3.6	2.6	0.8	0.5
10	3.1	1.6	0.8	0.3
11	2.4	1.7	0.8	0.7
12	1.9	1.1	0.8	0.6
Racial or ethnic group				
White	2.8	1.7	0.8	0.5
Black	2.0	1.3	1.0	0.6
Hispanic	3.2	2.6	0.7	0.5
Other[b]	1.8	1.6	1.0	1.0
Sex				
Female	1.8	1.0	0.5	0.3
Male	3.7	2.5	1.0	0.8
HIV instruction				
Yes	2.1	1.3	0.6	0.3
No	3.5	2.3	1.0	0.8

[a]Difference in sample size due to missing data from 8096 total.
[b]Includes Native Americans, Alaskan Natives, Asian or Pacific Islanders, and persons of other unspecified groups.

TABLE 5—Final Parameter Estimates for Regression of Ever Injecting Drugs on HIV Instruction, HIV Knowledge, and Demographic Characteristics, 1989 Secondary School Student Health Risk Survey

Variable[a]	b[b]	Standard Error of b
Intercept	−2.016	0.626
HIV instruction	0.183	0.288
HIV knowledge*	−0.145	0.041
Male*	1.037	0.342
Black	−0.014	0.356
HIV instruction × male*	−0.694	0.326
HIV instruction × Black*	−1.722	0.668

[a]Reference categories for HIV instruction, male, and Black are no HIV instruction, female, and other racial or ethnic group, respectively.
[b]Unstandardized regression coefficient.
*$P < .05$.

TABLE 6—Final Parameter Estimates for Regression of Injecting Drugs in the Last Year on HIV Instruction, HIV Knowledge, and Demographic Characteristics, 1989 Secondary School Student Health Risk Survey.

Variable[a]	b[b]	Standard Error of b
Intercept	−1.163	0.859
HIV instruction	0.039	0.421
HIV knowledge*	−0.187	0.048
Age	−0.164	0.092
Male*	1.212	0.423
HIV instruction × male	−0.663	0.512

[a]Reference categories for HIV instruction and male are no HIV instruction and female, respectively.
[b]Unstandardized regression coefficient.
*$P < .05$.

TABLE 7—Final Parameter Estimates for Regression of Ever Sharing Needles on HIV Instruction, HIV Knowledge, and Demographic Characteristics, 1989 Secondary School Student Health Risk Survey

Variable[a]	b[b]	Standard Error of b
Intercept	−2.549	1.179
HIV instruction	0.802	0.832
HIV knowledge*	−0.230	0.080
Male	1.440	0.837
HIV instruction × male	−1.638	0.850

[a]Reference categories for HIV instruction and male are no HIV instruction and female, respectively.
[b]Unstandardized regression coefficient.
*$P < .05$.

drugs when compared with males who had not received HIV instruction and with females. Further, Black students who had received HIV instruction were less likely to have ever injected drugs when compared with Blacks who had not received instruction and with students of other racial or ethnic groups. No significant direct effect of HIV instruction on ever injecting drugs was found.

Drug injection in the last year was significantly associated with level of HIV knowledge and sex (Table 6). The greater the level of knowledge, the less likely students were to have injected drugs in the last year, but, as with lifetime drug-injection behavior, males were significantly more likely than females to report the behavior in the last year. Again, there was no direct effect of HIV instruction on last year drug-injection behavior.

Similar results were found for those who reported ever sharing needles or sharing needles in the last year (Tables 7 and 8). The greater the level of knowledge, the less likely students were to have ever shared needles or to have shared needles in the last year. While the difference between males and females in ever sharing needles was not quite statistically significant, males were significantly more likely than females to have shared needles in the last year.

In additional analyses, not shown here, we found that students who reported injecting drugs and sharing needles were more likely to report multiple sexual partners (defined as two or more) and to report never using condoms when having sex.

TABLE 8—Final Parameter Estimates for Regression of Sharing Needles in the Last Year on HIV Instruction, HIV Knowledge, and Demographic Characteristics, 1989 Secondary School Student Health Risk Survey

Variable[a]	b[b]	Standard Error of b
Intercept	−1.040	0.790
HIV knowledge*	−0.371	0.063
Male*	0.950	0.435

[a]Reference category for male is female.
[b]Unstandardized regression coefficient.
*$P < .05$.

Also, in additional regression analyses, we found that within the subgroup of students who reported multiple sexual partners and within the subgroup of those who reported never using condoms, HIV instruction and knowledge had the same effect on ever injecting drugs and ever sharing needles that was demonstrated in our analysis of all students. The greater the level of knowledge, the less likely students were to report ever injecting drugs and ever sharing needles.

Furthermore, within the subgroup of students with multiple sexual partners, Black students who had been taught about HIV were less likely than Black students who had not received instruction and students of other racial or ethnic groups to report ever injecting drugs. Finally, within the subgroup of students who reported never using condoms, males who had received HIV instruction were less likely than noninstructed males and females to have ever injected drugs. Thus, even for students who engage in risky sexual behaviors, HIV knowledge and, for some students, HIV instruction continue to exert a positive influence on HIV-related drug behavior.

Finally, we examined the relationship between each HIV knowledge item and each of the drug behavior variables to see which items might be most predictive of the behaviors. Most highly negatively correlated with ever injecting drugs and ever sharing needles were correct responses to four items that asked about transmission of HIV through sharing needles and sexual intercourse.

Although HIV instruction did not have a significant direct effect on drug-injection and needle-sharing behavior independent of demographic characteristics, it was significantly associated with HIV knowledge (Table 9). When we controlled for demographic characteristics, students who had received HIV instruction were more likely than students who had not received instruction to have a greater level of HIV knowledge. Specifically, students who received HIV instruction in school answered, on average, almost one additional item correctly.

Age, being male, and racial or ethnic group were also significantly related to HIV knowledge. Students who were older and those who were White were more likely to have a higher level of knowledge about HIV, compared to younger students and those in the other racial or ethnic groups. Males, on the other hand, had a significantly lower level of knowledge about HIV than females did.

Discussion

Although the results of our survey showed that illicit drug-injection behavior was not common among the nation's high school students, it was reported by some students—more frequently by males and by those who had not received HIV instruction in school. The percentage of high school students who have injected illicit drugs is a cause for concern, however. If 2.7% of the nation's 9th- through 12th-grade students have injected drugs and 0.8 percent have shared needles, then approximately 345 100 and 102 200,[24] respectively, may have already engaged in behaviors that can quite efficiently transmit the virus that causes AIDS.

Moreover, students who inject drugs may be less likely to remain in school than students who do not. The observed decrease in drug-injection behavior by grade, and to some extent by age, could be attributed to the fact that dropout rates increase with grade—a phenomenon that, in turn, may be exacerbated by drug use. Indeed, one study showed a higher rate of drug use among dropouts than among in-school youth.[25]

When HIV instruction and demographic characteristics were taken into account, HIV knowledge was found to be associated with lower levels of drug-injection and needle-sharing behavior. And although males were more likely than females to report these behaviors, the results suggested that males who had received HIV instruction were less likely than uninstructed males to have ever injected drugs. Similarly, Blacks who had received HIV instruction were significantly less likely to report ever injecting drugs than were Black students who had not received instruction and students of other racial and ethnic groups. These results are particularly promising, because Blacks and males in the United States are disproportionately affected by AIDS.[26–28] Further, HIV instruction was significantly associated with an increased level of HIV knowledge among students, which in turn, was associated with lower levels of drug-injection and needle-sharing behavior.

Although these results have important implications for HIV education, some limitations of the data are important to consider. First, because the SSSHRS was a cross-sectional survey, causal relationships among the variables cannot be established. In addition, other variables that may have an effect on drug-injection behavior could not be accounted for in the analysis—for example, educational performance, socioeconomic status, peer norms, or drug availability. Also, because the data were self-reported, under- or overreporting may have occurred, especially given the sensitive nature of the behavior questions. Finally, since the data are school-based, the results can be generalized only to adolescents who attend school.

The data, however, were from a large, population-based sample; a number of variables were available for control; and more than one type of behavioral outcome was assessed. The prevalence of drug-injection behavior among secondary school students in the United States found by the SSSHRS was similar to other estimates of drug injection in this population,[12,15] and the analysis demonstrated that HIV knowledge does affect this behavior. In a study of adults, HIV knowledge was also shown to be a predictor of behavioral risk reduction.[29] Further, the effect of knowledge was significant for each indicator of drug-injection behavior in the current analysis. A similar effect of HIV knowledge was found when it was examined in relation to sexual behavior and condom use in the SSSHRS; that is, the greater the level of knowledge, the fewer the number of sexual partners and the higher the level of condom use.[9] Our results were also strengthened by an association between increased levels of HIV knowledge and lower levels of reported drug-injection and needle-sharing behavior within two subgroups of students at increased risk of HIV—those who reported two or more lifetime sexual partners and those who reported never using condoms when having sex.

TABLE 9—Final Parameter Estimates for Regression of HIV Knowledge on HIV Instruction and Demographic Characteristics, 1989 Secondary School Student Health Risk Survey

Variable[a]	b[b]	Standard Error of b
Intercept	12.613	0.298
HIV instruction[a]	0.953	0.091
Age[a]	0.110	0.040
Male[a]	−0.176	0.072
White[a]	1.047	0.277

[a]Reference categories for HIV instruction, male, and White are no HIV instruction, female, and other racial or ethnic group, respectively.
[b]Unstandardized regression coefficient.
[c]$P < .05$.

In our survey and in others, adolescents were quite knowledgeable about AIDS and HIV infection—at least about the major modes of HIV transmission.[10–15,17,18] But few previously reported studies focused on behavior, and only one that we found,[11] which was limited to a single high school, also examined determinants of HIV-related behaviors in youth.

While our analysis showed a relationship between knowledge and HIV-related drug behavior, we do not want to imply that knowledge is the only factor that may be related to behavioral change. Other factors, not measured in the current analysis, may be important predictors of behavioral change in addition to knowledge, for example, certain attitudes, social skills and support, perceived norms, and perceived efficacy in avoiding risks.[30] Thus, it should not be inferred from the results that educational programs designed to increase HIV knowledge alone will have an effect on behavior. But, while other factors that may affect drug-injection behavior need to be investigated, our results suggest that knowledge about HIV may be one factor in the complex process of behavioral change. Comprehensive school health education programs that increase knowledge, as well as specific skills that may influence HIV-related behavior, could be an important means to help prevent the further spread of HIV infection among our nation's youth. □

Acknowledgments

The material in this article was presented at the 6th International Conference on AIDS, San Francisco, Calif, June 1990; and at the 118th Annual Meeting of the American Public Health Association, New York, NY, September/October 1990.

The authors are grateful for the thoughtful comments of Dr. Richard Rubinson of Emory University, Atlanta, Ga, on earlier versions of this manuscript, and for the technical assistance of Dr. Donna Stroup of the Centers for Disease Control.

References

1. Manoff SB, Rogers MF, D'Angelo LJ, Dondaro TJ. The epidemiology of AIDS in adolescents and young adults. *J Adolesc Health Care.* 1987;8:307. Abstract.
2. Vermund SH, Hein K, Gayle HD, Cary JM, Thomas PA, Drucker E. Acquired immunodeficiency syndrome among adolescents: case surveillance profiles in New York City and the rest of the United States. *Am J Dis Child* 1989;143:1220–1225.
3. Munoz A, Wang M, Bass S, et al. Acquired immunodeficiency syndrome (AIDS)-free time after human immunodeficiency virus type 1 (HIV-1) seroconversion in homosexual men. *Am J Epidemiol.* 1989;130:530–539.
4. Centers for Disease Control. *HIV/AIDS Surveillance Report.* Atlanta, Ga: AIDS Program, Center for Infectious Diseases, CDC; July 1990.
5. Sonenstein FL, Pleck JH, Ku LC. Sexual activity, condom use, and AIDS awareness among adolescent males. *Fam Plann Perspect.* 1989;21:152–158.
6. Hofferth SL, Kahn JR, Baldwin W. Premarital sexual activity among U.S. teenage women over the past three decades. *Fam Plann Perspect.* 1987;19:46–53.
7. Zelnik M, Kantner JF. Sexual activity, contraceptive use, and pregnancy among metropolitan-area teenagers: 1971–1979. *Fam Plann Perspect.* 1980;12:230–237.
8. Zelnik M. Sexual activity among adolescents: perspective of a decade. In: McAnarey ER, ed. *Premature Adolescent Pregnancy and Parenthood.* New York, NY: Grune and Stratton; 1983.
9. Anderson JE, Kann L, Holtzman D, Arday S, Truman B, Kolbe LJ. HIV/AIDS knowledge and sexual behavior among high school students. *Fam Plann Perspect.* 1990;22:252–255.
10. Strunin L, Hingson R. Acquired immunodeficiency syndrome and adolescents: knowledge, beliefs, attitudes, and behaviors. *Pediatrics.* 1987;79:825–828.
11. Reuben N, Hein K, Drucker E, Bauman L, Lauby J, Silver E. Relationship of high-risk behaviors to AIDS knowledge in adolescent high school students. *J Adolesc Health Care.* 1988;9:261. Abstract.
12. Centers for Disease Control. HIV-related beliefs, knowledge, and behaviors among high school students. *MMWR.* 1988;37:717–721.
13. Centers for Disease Control. Results from the National Adolescent Student Health Survey. *MMWR.* 1989;38:147–150.
14. Hingson R, Strunin L, Berlin B. Acquired immunodeficiency syndrome transmission: changes in knowledge and behaviors among teenagers, Massachusetts Statewide Surveys, 1986 to 1988. *Pediatrics.* 1990;85:24–29.
15. Centers for Disease Control. HIV-related knowledge and behaviors among high school students. *MMWR.* 1990;39:1–7.
16. Becker MH, Joseph JG. AIDS and behavioral change to reduce risk: a review. *Am J Public Health.* 1988;78:394–410.
17. Price JH, Desmond S, Kukulka G. Perceptions and misperceptions of AIDS. *J Sch Health.* 1985;55:107–109.
18. DiClemente RJ, Zorn J, Temoshok L. Adolescents and AIDS: a survey of knowledge, attitudes, and beliefs about AIDS in San Francisco. *Am J Public Health.* 1986;76:1443–1445.
19. Greenberg RS, Kleinbaum DG. Mathematical modeling strategies for the analysis of epidemiologic research. *Annu Rev Public Health.* 1985;6:223–245.
20. SAS Institute Inc. SAS Users Guide: Statistics, Version 5 edition. Cary, NC: SAS Institute; 1985.
21. Shah BV. *SESUDAAN: Standard Errors Program for Computing of Standardized Rates from Sample Survey Data.* Research Triangle Park, NC. Research Triangle Institute; 1981.
22. Holt MM. *SURREGR: Standard Errors of Regression Coefficients from Sample Survey Data.* Research Triangle, NC:Research Triangle Institute; 1977.
23. Shah BV, Folsom RE, Harrell FE, Dillard CN. *RTILOGIT: Procedure for Logistic Regression on Survey Data.* Research Triangle Park, NC: Research Triangle Institute; 1987.
24. US Bureau of the Census. *Current Population Reports,* Series P-20, No. 443, School Enrollment—Social and Economic Characteristics of Students: October 1988 and 1987. Washington, DC: US Goverment Printing Office; 1990.
25. Chavez EL, Edwards R, Oetting ER. Mexican American and White American school dropouts' drug use, health status, and involvement in violence. *Public Health Rep.* 1989;104:594–604.
26. Centers for Disease Control. CDC Surveillance Summaries, July 1988. *MMWR.* 1988;37(no. SS-3):1–10.
27. Centers for Disease Control. AIDS and human immunodeficiency virus infection in the United States: 1988 update. *MMWR.* 1989;38(suppl. no. S-4).
28. Centers for Disease Control. Update: acquired immunodeficiency syndrome—United States, 1989. *MMWR.* 1990;39:81–86.
29. Emmons CA, Joseph JG, Kessler RC, Wortman CB, Montgomery SB, Ostrow DG. Psychosocial predictors of reported behavior change in homosexual men at risk for AIDS. *Health Educ Q.* 1986;13:331–345.
30. Kolbe LJ. Behavioral epidemiology of AIDS: an organizing model. In: Kulstad R, ed. *AIDS 1988: AAAS Symposia Papers.* Washington, DC: American Association for the Advancement of Science; 1988.

Needle Sharing in The Netherlands: An Ethnographic Analysis

Jean-Paul C. Grund, PhD, Charles D. Kaplan, PhD, and Nico F. P. Adriaans

ABSTRACT

Background. Needle sharing has been reported to be the main cause of the rapid spread of the human immunodeficiency virus (HIV) among injecting drug users. Risk behaviors such as needle sharing are, however, the end result of complicated interaction patterns in drug user networks, which have their specific rules and rituals, and larger social structures and official drug policy.

Methods. To study these interaction patterns we examined the drug administration rituals of heroin addicts in Rotterdam, The Netherlands. Intensive ethnographic descriptions were collected by participant observation.

Results. In less than 10% of the observed self-injections unsafe syringes were used. In 68% of the self-injections new, sterile syringes were used. Needle sharing as a planned sequence was not observed. Sharing was determined primarily by the availability of syringes, experience with the injecting ritual, and drug craving. In all observed needle-sharing events, subjects were aware of the risks involved and undertook efforts to clean the injection equipment.

Conclusions. In contrast to psychological approaches aimed at reducing individual "risk behavior," these findings suggest that HIV prevention can be made more effective if active drug injectors are organized to help themselves and their peers prevent high-risk exchange situations. (*Am J Public Health.* 1991;81:1602–1607)

Introduction

Injecting drug users (IDUs) are at risk for the human immunodeficiency virus (HIV) and other viral and microbacteriological infections through their use of contaminated syringes and needles,[1,2] generally termed "needle sharing." (Throughout the text, unless otherwise specified, the term needle sharing refers to the sharing of both needles and syringes.) In The Netherlands, only a minority of drug addicts inject. Most Dutch addicts smoke their heroin and cocaine from aluminum foil (chinesing or chasing the dragon).[3] Compared with neighboring countries and the United States, the availability and quality of heroin and cocaine on the Dutch illegal market have stabilized over the years at relatively high levels and these drugs may be purchased at moderate prices.[4] These economic factors were prerequisites for the diffusion[5,6] of the Asian practice of heroin smoking into the Dutch heroin-using population. However, the drug-injecting minority has not been overlooked by the Dutch government and helping organizations. Since the mid-1970s risk reduction and harm minimization strategies have been in place for all addicts.[7,8]

Few HIV seroprevalence studies have been conducted in The Netherlands. In a selected group that may not represent all drug users in Amsterdam, van den Hoek et al. found a HIV seroprevalence of 33% at entry into the study.[9] A seropositivity rate of 4.8% was found in a nonrepresentative sample from outside the large urban centers.[10] In a Rotterdam study on a sample of "extreme problematic drug users" in methadone maintenance, seropositivity was found to be 9.7% in 1986 and 6.5% in 1987.[8] As a comparison, among methadone clients in Rotterdam, self-reported lifetime prevalence for hepatitis is 19%, for gonorrhea 24%, and for syphillis 5%.[11]

Most of the current acquired immunodeficiency syndrome (AIDS) behavioral research consists of studies relating HIV/AIDS prevalence to HIV-risk factors, such as needle sharing. Few published studies describe what actually happens in such needle-sharing events. This article presents an ethnographic study of the patterns and circumstances of drug use and attempts to determine what variables are associated with risk behaviors such as needle sharing.

Methods

Data collected by self-report techniques may be highly biased if they reflect socially undesirable or criminalized behavior. Ethnographic field observations provide an opportunity to overcome the potential disadvantages of questionnaires or interviews by directly recording behaviors, thus reducing the biases of memory, self-perception, fear, and mistrust.[12] Nevertheless, the validity and generalizability of ethnographic study findings have been questioned.[13] The researcher's presence can alter the context and behaviors of the study group. However, the collective experience of the ethnography of socially undesirable behaviors indicates that with

The authors are with the EUR Addiction Research Institute (IVO) of the Erasmus University Rotterdam.

Requests for reprints should be sent to Jean-Paul C. Grund, EUR Addiction Research Institute (IVO), Medical and Health Sciences Faculty, Erasmus University Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands.

This paper was submitted to the journal February 27, 1990, and accepted with revisions April 23, 1991.

due care and time the researcher can become "part of the furniture."[13] Group members live within well-established traditions that structure and limit their actions. While the presence of the researcher may introduce a new constraint, the weight of tradition generally prevails.[14] The validity of ethnographic data can be further improved through appropriate controls on site and subject selection, observational strategies and protocols of data recording, and the development of trust between researchers and study subjects.[15]

The data to be presented were collected between February 1988 and April 1989 as part of an ongoing Rotterdam research project of the drug-taking rituals of heroin addicts. The principal methodology was "street" ethnography. Intense participant observation of addicts' self-administration of heroin and cocaine at dealing places, their homes, and public places was conducted. This observation was structured using an observational protocol of endogenous and exogenous cues pertaining to the heroin rituals of injecting and smoking.[16] The Appendix presents a summary table of the inventory of events indexed by the protocol.

Trust, acceptance, and credibility were developed by a field research team consisting of a community field worker (a respected ex-addict) and the principal investigator, both of whom were known to the participants prior to the study. During informal conversations with drug users, the project's goals were discussed. Other than providing an occasional sandwich or cup of coffee, participants received no fees. Enduring relationships have developed with some study participants who eventually became "key informants" and provided important supplemental information.

No structured interviews were held; when possible, additional information was collected from informal conversations. Detailed fieldnotes of 95 rituals were recorded. In 44 observations, subjects were smoking; in 2, subjects were snorting; and in 49, subjects were injecting.

The observations were recorded in the afternoon (94%) and early evening (6%) in two neighborhoods of Rotterdam characterized by high drug activity; 93% of the observations were recorded at 14 different houses where addicts live, deal drugs, or both. In 53% of the observations those houses were legally rented, in 9% they were subleased, and in 38% the house was squatted. The remaining observations were recorded in public places and abandoned buildings. All locations were initially selected through neighborhood exploration and key-informant information. At times, the research team accompanied participants on their daily rounds. In this way they were also introduced to new dealing places, using places, and private homes. During the fieldwork, some of the dealing places were closed down by the police. Although these "busts" caused some turmoil, they did not seem to have a significant impact on drug use. New places quickly opened up, sometimes on the same day at the same address. Often the researchers were introduced to the owner of a new address by a key informant. By following the dynamics of the drug scene, most of the neighborhood dealing sites could be observed.

Fieldnotes were produced independently by the researchers after each fieldwork session or the following morning, based on short notes taken during or immediately after the observations. The fieldnotes were processed using ETHNOGRAPH, a computer program for the analysis of qualitative data.[17] The data were coded in line with the observational protocol; the codes evolved constantly during the data collection and concurrent coding process. In order to complement this qualitative analysis, the data were quantified by counting significant events and by recording available demographic and background characteristics of the subjects. The numbers presented in this article were generated through these quantifications and are intended to support the qualitative analysis. Although this selected sample may not represent all drug users in Rotterdam, this is a common feature in most studies which examine "hidden" populations.

The total number of subjects contacted was 192 (168 males and 24 females). By comparing independent field-worker estimates, ages of 106 subjects could be specified: 30% were 25 years or younger, 50% were between 25 and 35, and 20% were over 35. This distribution resembles that found in the Rotterdam registration system of heroin addicts in methadone treatment (RODIS).[11] For 23% of the subjects injecting was the main mode of drug administration, while the remaining 77% smoked; 96% used both heroin and cocaine. The field work established that both heroin and cocaine were available at all dealing addresses. That cocaine has become increasingly important to Dutch heroin addicts is verified by the RODIS system, in which 72% of the treatment population uses cocaine.[11]

TABLE 1—Status and Origin of Syringes in Observed Self-Injections (N = 44)

Status and Origin	n	%
New		
In possession of subject	9	21
Received at place of injection	21	48
Used		
In possession of subject	10	23
Received/found at place of injection	2	5
Received from other intravenous drug user	2	5

Results

Sharing behaviors were found to be frequent and important events for the drug users observed in this study. The sharing of valued things such as housing, food, and clothing is an everyday occurrence tied to survival needs in extreme circumstances. The sharing of drugs fits into this wider pattern of daily interaction and exchange[18] and has been observed for both IDUs and non-IDUs. Among IDUs this interaction pattern may have important implications for HIV transmission.[19] The sharing of drug paraphernalia was also observed: at some addresses spoons, water containers, and lemon juice were available for common usage, and some drug users shared their knives or lighters upon request. However, needle sharing as a planned or stereotyped sequence in which two or more people share, one after the other, the only available syringe was never observed. Most drug users seemed aware of the risks and consequences of needle sharing.

In 68% of self-injection events, a new syringe was used. In 23% IDUs reused their own syringe. In less than 10% of the self-injections, a potentially unsafe syringe was reused (Table 1). Used syringes that were found or received from others were not adequately cleaned (i.e., cleaned with water only). The unsafe self-injections were recorded at addresses where injecting was permitted and in a public place. Because most houses don't allow injecting, IDUs used their drugs in private settings more often than smokers. These house rules may prevent needle sharing. However, some IDUs let their friends and acquaintances inject at their place, sometimes in return for a "taste" of the drug. At some of these using places and at some of the dealing places that allow people to inject, clean syringes supplied by an out-

reach and needle exchange program[20] were dispensed as illustrated in the following fieldnote.

> **1. "Is that an old spike you want to use?"** Although it isn't stated directly, Karel agrees on Jerry taking a shot at his place. Jerry wants to shoot up pure cocaine. He puts his spike on the table and asks Karel for a spoon. Karel asks: "Is that an old spike you want to use?" Jerry: "Well, old, I've used it one time before, so it's still good to use." Karel: "I've got some new ones left from the exchange." He gives a syringe to Jerry asking "You want some more for tonight or the weekend?" Jerry: "If you can spare them I'll take some with me." Karel gives him four syringes.

For many IDUs a strong incentive exists to frequent those places where clean (and sharp) syringes are freely available. Nonetheless, despite the availability of syringes, there were times when we observed IDUs in situations of heightened needle-sharing risk. In those cases, traditional subcultural sharing norms conflict with the more recent norms regarding safe use. The following fieldnote recorded at the "shooting room" of a dealing address documents such a situation.

> **2. "Don't be afraid, I've got no AIDS."** From the corner where she's still busy with the needle in her hand Anja asks: "Do you have a new syringe for me?" Eric answers, "No, this is the only one I got. I was lucky, one minute before the pharmacy closed I bought it." Then Anja asks: "Can't I use it when you're finished. I can't use these anymore," pointing at the syringes laying around her on the floor. "They're all blunt, I can't hit a vein," she says as if she's crying or starting to cry. "Or let me only use the needle then. Please let me, I will clean it for you. Don't be afraid. I've got no AIDS. I've been tested recently at the methadone program." Eric still refuses: "I would like to help you but it's the only one I've got. I never lend out my spike to someone else, nobody, not when I have to use it myself again. When I'm not using it again it's their own risk."

Although needle sharing did not take place, the fieldnote illustrates the often emotional pressure to share a syringe. In this case, in spite of the pressure Eric decided in the interest of his personal health and against the norm of sharing. Nevertheless, the situation can often be tense.

When comparing fieldnotes 1 and 2 the functioning of a variable of perceived responsibility for safe use becomes apparent. In fieldnote 1, the perceived responsibility is given a collective quality. Because the needle exchange program dispenses supplies to IDUs, the possession of ample syringes allows users to reinforce the subcultural norm of sharing, yet in a safe way. In contrast, in fieldnote 2, the perceived responsibility is marked by a strong quality of the "rugged individual junkie" who resists the pressure of community sharing norms by adhering to a personal behavior code.

Although both collective and individual norms of responsibility operate to minimize and manage the risks of needle sharing, there are still situations in which unsafe needle behaviors can be observed. Fieldnote 3, for example, was recorded on the first floor of a house occupied by squatters. The first and third floors of the three-story house were in use. It did not have running water; water was carried to the house in containers from a garage next door. The house was inhabited by a group of older IDUs. At times the group offered shelter to other drug users. They also allowed other drug users to deal heroin, cocaine, or amphetamine in exchange for money or drugs. Occasionally they themselves dealt drugs as was the case in fieldnote 3. This house also was supplied with sterile syringes by the outreach program. However, at the time of observation it was unclear whether there were any new syringes left. Jack (the doorman) let in Billy and Dirk. They all knew one another. Meanwhile some other users went in and out.

> **3. "I don't give a shit about AIDS."** Billy asks Dirk what he wants. "Let's do coke first and then a cocktail," Dirk replies. Billy has a syringe wrapped in aluminum foil. He does not want to wait for a new syringe. Dirk does not have one with him and starts searching. He asks Jack if there are any new ones left. "I don't know," Jack replies. "Maybe upstairs. Ask Karel, he's there." They call Karel several times but he does not answer. Then Dirk finds a syringe in a cupboard. It is unclear who it belongs to. He rinses it with water. He pulls up water twice from a cup he has filled from the water container and squeezes it through the needle. Billy mumbles something about AIDS. Dirk says "Ik heb schijt aan AIDS." (I don't give a shit about AIDS).

Dirk used this syringe two times. First he shared a dose of cocaine with Billy and within 30 minutes he shared a "cocktail" (heroin–cocaine mixture) with Billy and Jack.

In fieldnote 4, two IDUs (Eric and Anja) were in the "shooting room" of a dealing address where injecting is allowed when Leo entered the room. Leo did not come to the address to buy drugs. Eric owed him money and he had heard that Eric was at the address.

> **4. "You have to clean it well."** "I wanted to ask you if you can pay back or otherwise if you could help me with a shot." Leo said. Eric was not able to pay Leo back but offered a cocktail. Leo gladly accepted: "Great man, you don't know how wonderful that is. I'm so glad I didn't miss you here." However, Leo was not in possession of a syringe. Leo asks Eric for a syringe: "I couldn't get a new one. The needle exchange at the Central Station had closed already." Eric tells Leo he only has his own, which he is not willing to share. Then Leo asks Anja if he can take one of her used syringes that are lying in front of her. Anja: "That's useless, they're all blunt, but if you want to try that's okay with me." She picks up several syringes from the floor and looks closely at the needle, comparing one with the other. Finally she makes her decision which one to give to Leo and gives it saying: "You have to clean it well." Leo goes with the syringe toward the sink and cleans it seven or eight times with cold water. To clean the plastic part of the needle, he moves it in such a way that there is some space between syringe and needle. He presses the plunger strongly so the plastic is cleaned under pressure. The water now does not come through the needle but shoots away through the little space between syringe and the plastic needle holder. Leo states: "It must be clean now." Anja tells him: "Don't worry, I'm checked for AIDS recently. I told Eric too."

In fieldnote 5 Mohammed and Abdul had obtained new syringes from the exchange program near the Central Railroad Station before they went to an address to buy cocaine and heroin. They were not allowed to inject at the address, so they went to a small greenhouse in a park. Mohammed prepared the jointly bought drugs. He then divided them by "front-loading," a process in which the solution is first drawn from the spoon into one syringe and then half of the solution is inserted into the hub of the second syringe.[19] Abdul wanted to check to see if the solution was equally divided between both syringes.

> **5. "Don't you trust me?"** Mohammed gives both syringes to Abdul and asks him: "Don't you trust me?" Abdul doesn't answer. He holds the syringes next to each other and stares at them. While doing this he accidentally drops a syringe. The needle falls straight into the ground. Abdul curses and so does Mohammed. Mohammed says, "Now you see what happens. Why don't you believe me?" Abdul picks up the syringe and looks closely at the needle. He asks Mohammed if he still can use it. Mohammed takes the syringe and runs the needle tip over his thumbnail. "No," he says, "there is a burr on it. It's not sharp anymore and it's dirty. You've got to get a new one." Abdul:

"No, I don't go back. Give me yours." Mohammed: "Then you have to wait until I'm ready." After Mohammed has taken his shot he starts cleaning his syringe with the water from the bottle. He puts some lemon juice in the cooker again, pulls it into his syringe, and shakes it so that it mixes with the water. He puts the needle back on it, holds it with two fingers, and presses the water–lemon solution through it as hard as possible. Mohammed explains: "The lemon bites and cleans the needle better." When he's ready, he puts the needle on Abdul's syringe.

Examination of fieldnotes 3, 4, and 5 indicates that needle sharing often is the result of complex and multiple factors. In all the fieldnotes, use of another's syringe was not planned. Rather, an unexpected situation occurred. In fieldnote 3, Billy and Dirk select the particular address because injection is allowed and new syringes are available. However, when they were ready to inject, new syringes were not available. In fieldnote 4, Leo did not have the money to buy drugs but learned by coincidence that Eric, who owed him money, was at the address. He did not have a syringe because the needle exchange had closed for the day. In fieldnote 5, although sufficient prevention measures were taken in advance, an unfortunate accident put one of the users at risk.

In these three fieldnotes an unanticipated change in the situation put the IDUs in the uncomfortable position of choosing between postponing or abstaining from a shot and an unsafe injection. They all chose the unsafe injection although they were well aware of the potential risks of their behavior as evidenced by their rather intensive efforts to clean the used syringes. In fieldnotes 3 and 4, AIDS is associatively mentioned while cleaning the syringes; in fieldnote 5, the use of lemon juice is presented as a cleaning method superior to using water only. However, the effectiveness of AIDS-related knowledge as a protective factor seems to depend on certain specific situational factors.

The significance of one of these situational factors, the intensity of drug craving, deserves special attention. Shortly before an injection, IDUs often become highly aroused. This arousal leads to preoccupation with the sequence that relieves craving. For some IDUs, this sequence begins when the drugs are obtained. For others, the preparation of the injection is the starting point. As one IDU states: "As soon as I put it on the spoon my stomach turns around and I know it's gonna happen, I'm gonna feel that intense rush." In all of the cases presented, postponing the injection would have caused noticeable stress. In fieldnote 3, Billy and Dirk were in high anticipation of an injection. They had already visited the address 30 minutes earlier, expecting to find a dealer working. However, the dealer had just left, so they went to another to buy. When they returned, Billy complained about the trouble they had finding a dealer. Dirk did not bother to see if there were any new syringes upstairs. By this time, the drug craving had become too intense for further deferral. In fieldnote 4, Leo explicitly expressed relief when he was offered an injection. In fieldnote 5, the accident happened only seconds before actual injection. Obviously, Abdul's craving had become so intense that he ignored Mohammed's advice to get another syringe. Furthermore, in all cases cocaine or a mixture of cocaine and heroin was injected. The addition of cocaine to the daily drug-using rituals has been observed in the field to be associated with an intensification of craving and a disruption of stabilized heroin-methadone patterns.

While situational factors play the most important role in needle sharing, certain structural factors can also be seen. One such factor is that of socially learned experience with the injecting ritual, internalized during the IDU career. This experience includes, among other things, protective skills that support safe needle use. The interaction between Billy and Dirk in fieldnote 3 provides an illustration of this factor. Billy, an experienced IDU, went to inject at a place where clean syringes are normally available. Nevertheless, he carried his own used syringe anticipating the absence of new ones. In contrast, Dirk, a novice IDU with no visible needle marks, did not bring a syringe. Immediately after injecting cocaine, he began smoking heroin. He identified himself, not as an injector, but as a smoker, stating: "I'm only shooting now and then, strictly speaking I am a chineser." Irrespective of his AIDS knowledge, Dirk's particular self-deception works against taking the appropriate precautions. His self-perception as a chineser provides a false sense of security that, in turn, leads to blasé attitudes such as "I don't give a shit about AIDS." In fieldnote 5, another case of inadequate socialization is found. Both users were still at an early point in their injecting careers. Mohammed, who reported he had been injecting for about half a year, is instructing the even less experienced Abdul, who was about to take his 10th lifetime injection.

Discussion

Increased availability of drugs is thought to affect consumption by increasing the prevalence of use.[21] Less attention has been given to the harmful effect of decreased availability of drugs. Availability-related variables, such as rising prices, decreasing purity, and unstable supplies can be seen as factors determining the onset of injecting.[22–24] The Dutch experience shows that if availability variables are relatively stable over time, minimizing economic pressure to initiate and maintain injecting, a predominant smoking pattern can develop.[4,25] Thus, injecting can be seen as adaptation to the conditions of decreased drug availability.

The availability of needles also plays a major role in needle sharing.[23,26] Syringes have traditionally been easy to obtain in The Netherlands. Since the AIDS epidemic began, the availability of syringes has increased even more due to the needle exchange programs. Furthermore, in The Netherlands, in contrast to other countries, the possession of injecting equipment has never been illegal.[23,26–28] Risk of arrest discourages IDUs from carrying their personal injection equipment, which limits on-the-spot availability and increases the frequency of sharing drug paraphernalia.

The results of this field study support the hypothesis that when drug and syringe availability is stable and possession of injection equipment is legal, needle sharing decreases markedly. Nevertheless, research in Europe and the United States documents that IDUs are indeed changing their behavior to less risky injecting practices, despite the absence of "Dutch" conditions.[29–31] There seems to be a growing awareness of health risks and a willingness to use drugs in safer, more responsible ways. Our field research found only limited incidence of irresponsible behavior in inexperienced IDUs or those experiencing craving intensified by cocaine. Several recent studies have shown a relationship between cocaine use, risk behaviors, and HIV-serostatus.[32,33] Paralleling other studies, our results show that, even under optimal conditions, IDUs can and do get into situations in which sterile injection equipment is not available. In contrast with novice IDUs, experienced injectors are likely to be more competent in managing such situations.[29,34]

In numerous studies, the social organization of the drug subculture has also been associated with needle sharing. Drug users are often organized in small friendship groups.[35,36] These friendship groups

are often linked in networks whose main activity is to obtain and distribute money and drugs,[35,36] which are often shared and used along with other necessities of life. Sharing and the reciprocal aid it promotes provide a practical and emotional balance to the daily hardship of addict life. In the pre-AIDS era, needle sharing fitted snugly into this pattern. Helping a fellow addict with a syringe was an expression of the almost universal subcultural code of "share what you have."[37] These sharing behaviors functioned not only to satisfy individual craving, but also to support the maintenance of the network through the expression of community solidarity and the instrument of economic exchange.[35] In the AIDS era, needle sharing no longer serves this purpose. Rather, it has been transformed into a threat to the individual drug user, the friendship groups and networks, and the drug subculture as a whole. As with most cultural shifts, the process is gradual and never complete. Residues of the traditional code still remain and can be observed in emotional appeals and convenient lapses in newly acquired knowledge.

In conclusion, our findings suggest that needle sharing often is the outcome of structurally or situationally determined social interaction. Knowledge alone is not enough to counter the pressure of social interaction and drug craving. In these stressful situations, if clean syringes are not available, the strength of addiction will ultimately lead IDUs to unsafe injection practices. While easy access and a sufficient supply of clean syringes are effective in reducing unsafe practices, as van den Hoek et al. conclude,[9] they are not enough. They recommend intensive counseling in future preventive education efforts. Our findings support their suggestion by identifying a number of factors that determine needle sharing. These factors should be addressed in counseling that focuses on the practical skills of safe drug use. Furthermore, our results indicate that changes in the social environment may be more important than changes in individual risk behaviors. Prevention efforts may be made more effective by mobilizing social resources directed at preventing risk situations. IDUs and their networks should have a prominent role in such approaches, an idea that is getting more attention.[38–40] In Rotterdam our findings have shown that outreach programs working with IDUs to reinforce positive protective factors such as rules of safe use, while also distributing syringes to unknown IDUs via known ones can be effective in changing the environment.[20] Utilizing the knowledge of drug users and their information and exchange networks to promote risk reduction through peer education and peer support might offer more perspective on a lasting behavior change than any other prevention effort. □

Acknowledgments

This research was supported by grant No. 900-556-039 from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Dutch Organisation for Scientific Research), Medical Science Field, Social Psychiatry Working Group. Opinions expressed in the article do not necessarily reflect the policy of the supporting organization.

We thank the drug users who participated in this study for their cooperation; Dr. Peter Blanken, EUR Addiction Research Institute, for his valuable suggestions and constructive review of an earlier draft version; and Dr. Jaap Toet, Rotterdam Municipal Health Service, Epidemiology Department, The Netherlands, for RODIS data and additional analysis on our request. We also thank L. Synn Stern, Montefiore Medical Center, Bronx, New York/DV8 Consultants for her help in translating this manuscript into English.

References

1. Brettle RP. Epidemic of AIDS related virus infection among intravenous drug abusers. *Br Med J.* 1986;292:1671.
2. Chaisson RE, Moss AR, Onishi R, Osmond D, Carlson JR. Human immunodeficiency virus infection in heterosexual intravenous drug users in San Francisco. *Am J Public Health.* 1987;77:169–172.
3. Buning EC, Coutinho RA, Brussel GHA van, Santan GW van, Zadelhoff AW van. Preventing AIDS in drug addicts in Amsterdam. *Lancet.* 1986;2:1435.
4. Kaplan CD, Janse HJ, Thuyns H. Heroin smoking in The Netherlands. In: *Drug Abuse Trends and Research Issues, Community Epidemiology Work Group Proceedings.* Rockville, Md: National Institute on Drug Abuse; December 1986. III-35–45.
5. O'Donnell JA, Jones JP. Diffusion of the intravenous technique among narcotic addicts in the United States. *J Health Soc Behav.* 1968;9:120–130.
6. Katz E, Levin ML, Hamilton H. Traditions of research on the diffusions of innovation. *Am Sociol Rev.* 1963;28:237–252.
7. Engelsman EL. Dutch policy on the management of drug related problems. *Br J Addict.* 1989;84:211–218.
8. Barends W. Routinematig HIV-onderzoek in een Rotterdams methadonprogramma. Medisch Contact 1988;43(2):58–60. (Routine HIV-testing in a Rotterdam methadone maintenance program)
9. Hoek JAR van den, Haastrecht HJA, Coutinho RA. Risk reduction among intravenous drug users in Amsterdam under the influence of AIDS. *Am J Public Health.* 1989;79:1355–1357.
10. Limbeek J van, Wouters L, Hekker AC, Cramer A. Een pilot-studie naar het voorkomen van personen met HIV-antistoffen in hulpverleningsprogramma's voor drugsverslaafden buiten de Randstad. Bilthoven, The Netherlands FZA, 1987. (A pilot study into the occurrence of HIV seropositive persons in drug treatment programs outside the major cities)
11. Toet J, Ven APM van de. Het RODIS uit de steigers: resultaten 1988 Rotterdam. Report 65, GGD Rotterdam afdeling Epidemiologie, 1989. (The Rotterdam Drug Information System: Results 1988)
12. Lambert EY, ed. *The Collection and Interpretation of Data from Hidden Populations.* NIDA research monograph 98. Rockville, Md: National Institute on Drug Abuse; 1990.
13. Agar MH. *Speaking of Ethnography.* Beverly Hills, Calif; Sage; 1986. Sage University Paper series on Qualitative Research Methods, Volume 2.
14. Becker H. In: *Sociological Work: Method and Substance.* New Brunswick, NJ: Transaction Books; 1970.
15. Douglas JD. Investigative social research. In: *Individual and Team Field Research.* Beverly Hills, Calif: Sage; 1976.
16. Grund J-PC, Kaplan CD, Adriaans NFP, Harms J. Heroin rituals: observational protocol of endogenous and exogenous cues. Rotterdam, The Netherlands: Addiction Research Institute (IVO), Erasmus University Rotterdam; 1988.
17. Seidel JV. The ETHNOGRAPH, Version 3.0 [computer program]. Littleton, Colo: Qualis Research Associates; 1988.
18. Grund J-PC, Kaplan CD, Adriaans NFP, Blanken P. Drug sharing and HIV transmission risks: the practice of "frontloading" in the Dutch injecting drug user population. *J Psychoactive Drugs.* 1991;23:1–10.
19. Grund J-PC, Kaplan CD, Adriaans NFP, Blanken P, Huisman J. The limitations of the concept of needle sharing: the practice of frontloading. *AIDS.* 1990;4:702–703.
20. Grund J-PC, Blanken P, Adriaans NFP, Kaplan CD, Barendregt C, Meeuwsen M. Reaching the unreached: an outreach model for "on the spot" AIDS prevention among active, out of treatment drug addicts. In: O'Hare PA, Newcombe R, Buning EC, Drucker E, and Matthews A, eds. *Reducing Drug Related Harm.* London, England: Routledge Press (in press).
21. Goldstein A, Kalant H. Drug policy: striking the right balance. *Science.* 1990; 249:1513–1521.
22. Casriel C, Rockwell R, Stepherson B. Heroin sniffers: between two worlds. *J Psychoactive Drugs.* 1988;20(4):37–40.
23. Power RM. The influence of AIDS upon patterns of intravenous use—syringe and needle sharing—among illicit drug users in Britain. In: Battjes RJ, Pickins RW, eds. *Needle sharing among Intravenous Drug Abusers: National and International Perspectives.* Rockville, Md: National Institute on Drug Abuse; 1988:75–88.
24. Parker H, Bakx K, Newcombe R., eds. *Living with Heroin: The Impact of a Drug 'Epidemic' on an English Community.* Philadelphia, Pa: Open University Press, Milton Keynes; 1988.
25. Grund J-PC, Adriaans NFP, Kaplan CD. Changing cocaine smoking rituals in the Dutch heroin addict population. *Br J Addict.* 1991;86:439–44.
26. Feldman HW, Biernacki P. The ethnography of needle sharing among intravenous drug users and implications for public pol-

icies and intervention strategies. In: Battjes RJ, Pickins RW, eds. *Needle Sharing among Intravenous Drug Abusers: National and International Perspectives.* Rockville, Md: National Institute on Drug Abuse; 1988:28–39.
27. Olievenstein C. Drug addiction and AIDS in France in 1987. In: Battjes RJ, Pickins RW, eds. *Needle Sharing among Intravenous Drug Abusers: National and International Perspectives.* Rockville, Md: National Institute on Drug Abuse; 1988:114–118.
28. Pascal CB. Intravenous drug abuse and AIDS transmission: federal and state laws regulating needle availability. In: Battjes RJ, Pickins RW, eds. *Needle Sharing Among Intravenous Drug Abusers: National and International Perspectives.* Rockville, Md: National Institute on Drug Abuse; 1988:119–136.
29. Stimson GV, Alldritt LJ, Dolan KA, Donaghoe MC, Lart RA. *Injecting Equipment Exchange Schemes: Final Report.* London, England: Monitoring Research Group; 1988.
30. Kokkevi A, Alevizou S, Stefanis C. AIDS related behavior and attitudes among IV drug users in Greece. Presented at the Fifth International Conference on AIDS, Montreal, Quebec, Canada, 1989. Abstract THDP 72.
31. Guydish J, Abramowitz A, Woods W, Newmeyer J. Sharing needles: risk reduction among IDU's in San Francisco. Presented at the Fifth International Conference on AIDS, Montreal, Quebec, Canada 1989. Abstract THDP 34.
32. Wiebel W, Ouellet L, Guydan C, Samairat N. Cocaine injection as a predictor of HIV risk behavior. Poster presented at the Sixth International Conference on AIDS, San Francisco, Calif, June 1990. Abstract FC 767.
33. Schoenbaum EE, Hartel D, Friedland GH. Crack use predicts incident HIV seroconversion. Paper presented at the Sixth International Conference on AIDS, San Francisco, Calif. June 1990. Abstract Th.C.103.
34. Schilling R, El-Bassel N, Schinke S, Botvin G. Risk behavior and attitudes among recovering IV drug users. Paper presented at the Fifth International Conference on AIDS, Montreal, Quebec, Canada, 1989. Abstract THDP 42.
35. Des Jarlais DC, Friedman SR, Sotheran JL, Stoneburger R. The sharing of drug injection equipment and the AIDS epidemic in New York City: the first decade. In: Battjes RJ, Pickins RW, eds. *Needle Sharing Among Intravenous Drug Abusers: National and International Perspectives.* Rockville, Md: National Institute on Drug Abuse; 1988:160–175.
36. Preble E, Casey JJ. Taking care of business—the heroin user's life on the street. *Int J Addict.* 1969;1:1–24.
37. Wieder DL. Telling the code. In: Turner R, ed. *Ethnomethodology: Selected Readings.* Middlesex, England: Penguin Education; 1974:144–172.
38. Grund J-PC, Stern LS, Kaplan CD, Adriaans FP, Drucker E. Drug use contexts and HIV consequences: the effect of drug policy on patterns of everyday drug use in Rotterdam and the Bronx. *Br J Addict.* (Special issue on AIDS).
39. Des Jarlais DC, Friedman SR. Shooting galleries and AIDS: infection probabilities and "tough" policies. *Am J Public Health.* 1990;80:142–144.
40. Chitwood DD, McCoy CB, Inciardy JA, et al. HIV seropositivity of needles from shooting galleries in South Florida. *Am J Public Health.* 1990;80:150–152.
41. Du Toit BM, ed. *Drugs, Rituals, and Altered States of Consciousness.* Rotterdam, The Netherlands: Balkema; 1977.
42. Agar MH. Into that whole ritual thing: ritualistic drug use among urban American heroin addicts. In: Du Tolt BM, ed. *Drugs, Rituals, and Altered States of Consciousness.* Rotterdam, The Netherlands: Balkema; 1977:137–148.

APPENDIX[16]

The observational protocol used in this investigation provides an instrument to guide field research of heroin routes of administration in naturalistic settings. For theoretical purposes, we have termed the routes of administration for heroin observed in naturalistic settings as heroin rituals. The concept of ritual has been used in previous studies of drug use.[41] Our operational definition of ritual has been adopted from M. H. Agar[42]: "For an event to be a ritual event, it must prescribe a sequence of psychomotor acts . . . [and] the prescribed psychomotor sequence must be invested with a special meaning for the person performing that sequence." (p. 141). The instrument is partitioned across two dimensions. The first dimension, heroin rituals, consists of two categories; spuiten (injecting) and chinezen ("chasing the dragon," smoking). The second dimension refers to observational cues, classified as exogenous and endogenous. Exogenous cues index external ecological events that correspond to the "prescribed psychomotor sequence." In addition, participant background characteristics, paraphernalia, and environmental features partly compose this category. Endogenous cues index internal psychosocial events that correspond to the "special meanings" of the rituals. Internal experience is routinely expressed and displayed through language or gesture thereby available for direct observation in naturalistic settings.

Table A provides an overview of endogenous and exogenous cues cross-classified by the rituals. The table has the character of "fuzzy (sub)sets" with overlap of the formal items assigned to its four cells. In addition, this classification may not be exhaustive: future research may add new dimensions, categories, or items to the instrument.

TABLE A—Summary Inventory of Observational Cues by Heroin Rituals

Cues	Rituals	
	Smoking	Injecting
Exogenous	Background characteristics	Background characteristics
	Argot	Argot
	Tools	Tools
	Drug-taking actions	Drug-taking actions
	Order	Order
	Prices	Prices
	Places	Places
	Time	Time
	Behavior	Behavior
	Rules	Rules
	Sanctions	Sanctions
	Relations	Relations
	Hierarchy	Hierarchy
	Hygiene	Hygiene
		External circumstances
Endogenous	Argot (jargon)	Argot (jargon)
	Glosses (idiom)	Glosses (idiom)
	Drug effects	Drug effects
	Emotions	Emotions

Socioeconomic Status and Electrolyte Intake in Black Adults: The Pitt County Study

ABSTRACT

Background. Although the inverse association between socioeconomic status (SES) and blood pressure has often been observed, little is known about the relationship between SES and dietary risk factors for elevated blood pressure. Therefore, this study described the distribution of dietary intakes of sodium, potassium, and calcium and examined the association between electrolyte intake and SES among 1784 Black men and women aged 25 to 50 residing in eastern North Carolina.

Methods. Household interviews were conducted in 1988 to obtain information on psychosocial and dietary correlates of blood pressure. Electrolyte intake (mg/day) was assessed using a food frequency questionnaire adapted to reflect regional and ethnic food preferences. SES was categorized into three levels defined by the participant's educational level and occupation.

Results. After adjustment for age and energy intake, potassium and calcium intake increased with increasing SES for both sexes. Sodium intake was high for all groups and did not vary markedly with SES, but sodium to potassium and sodium to calcium ratios decreased with increasing SES. In addition, high SES individuals were more likely to believe that diet affects risk for disease and to report less salt use at the table and less current sodium consumption than in the past.

Conclusion. These data indicate that nutritional beliefs as well as the consumption of electrolytes are associated with SES in Black adults. (*Am J Public Health.* 1991;81:1608–1612.)

Ann M. Gerber, PhD, Sherman A. James, PhD, Alice S. Ammerman, DrPH, Nora L. Keenan, MPH, Joanne M. Garrett, PhD, David S. Strogatz, PhD, Pamela S. Haines, DrPH

Introduction

As noted in several recent reviews,[1–3] socioeconomic status (SES) is inversely associated with mean blood pressure and the prevalence of hypertension. This inverse association has been observed among Blacks in the initial screening of the Hypertension Detection and Follow-up Program[4] and in a number of smaller, geographically dispersed community studies.[5–9] It is not yet clear what combination of biological, psychosocial, or environmental factors contributes to this relationship. Diet may be among the important environmental factors that influence the inverse relationship between SES and blood pressure.

The objectives of this report are to describe the distribution of dietary intakes of sodium, potassium, and calcium among Black men and women and to determine whether intake of these nutrients varies by SES. This report is part of the Pitt County Study,[10] which was designed to investigate psychosocial and dietary influences on the development of hypertension in young and early middle-aged Black adults residing in an eastern North Carolina community. This community is part of a larger area within the southeastern United States known to have high mortality from stroke and other hypertension-related disorders.[11]

A variety of nutritional factors, including intake of sodium, potassium, and calcium, have been associated with blood pressure level,[12–15] although intracommunity comparisons[16] are, in general, less likely to demonstrate significant effects than studies that examine diverse, multicountry experiences.[17,18] While most studies have not explicitly investigated differences in electrolyte intake by SES, variations in food consumption patterns or macronutrient intake by educational level and income have been reported.[19–23] Increased incomes are associated with greater consumption of meats, fruits and vegetables, and dairy products.[20,23] This is consistent with data from the Second National Health and Nutrition Examination Survey (NHANES II), which indicated that young adults below the poverty level consumed less potassium and calcium than those above the poverty level.[24] Sodium intake per 1000 calories did not differ by poverty status.[24]

The nutrient intake–economic status relationships investigated in these national surveys[23–26] and community-based populations[19] were not consistently stratified by race. However, large sample surveys of nutritional status and dietary behavior in the United States have shown lower

Ann M. Gerber, Sherman A. James, Nora L. Keenan, and David S. Strogatz are with the Department of Epidemiology, School of Public Health; Alice S. Ammerman and Pamela S. Haines are with the Department of Nutrition, School of Public Health; and Joanne M. Garret is with the Health Services Research Center—all at the University of North Carolina at Chapel Hill. Sherman A. James is also with the Department of Epidemiology, School of Public Health and Survey Research Center, Institute for Social Research, University of Michigan. David S. Strogatz is also with the School of Public Health, State University of New York at Albany and the New York State Department of Health, Albany, NY.

Requests for reprints should be sent to Sherman A. James, PhD, University of Michigan, Department of Epidemiology, School of Public Health, 109 Observatory Street, Ann Arbor, MI 48109.

This paper was submitted to the journal September 28, 1990, and accepted with revisions June 18, 1991.

intakes of potassium, calcium, and magnesium in Blacks compared to Whites. Differences in calcium intake between Blacks and Whites may be due to a higher overall prevalence of lactase deficiency among Blacks in the United States.[27] In addition, Langford et al.[28] reviewed the diet–blood pressure relationship and concluded that the lower dietary intake of potassium and possibly calcium among US Blacks may be a major factor in explaining their well-known excess risk for hypertension.[1] Little is known about whether these nutritional determinants of blood pressure vary by SES among Blacks. This report is intended to help clarify the latter issue and, in doing so, facilitate more careful planning of dietary interventions to reduce risk for hypertension in US Blacks.

TABLE 1—Composite Measure of Socioeconomic Status (SES): The Pitt County, NC, Study

Education	Current or Last Occupation: Hollingshead 1–2 (Unskilled)	Hollingshead 3 (Semiskilled)	Hollingshead 4–9 (Skilled or white collar)
< High school	SES1 (n = 297)	SES1 (n = 158)	SES2 (n = 49)
High school grad	SES1 (n = 217)	SES2 (n = 274)	SES3 (n = 189)
> High school	SES2 (n = 63)	SES2 (n = 117)	SES3 (n = 316)

Methods

Study Population

The study population is composed of a community-based sample of 25- to 50-year-old Black men and women residing in Pitt County, NC, in 1988. The sampling design identified housing units through an area probability sample. Area segments of housing units were selected with probabilities proportional to the expected number of Blacks residing in each area. All age-eligible men and women within selected housing units were invited to participate in the study. A major sampling objective was to obtain an economically heterogeneous sample. Eighty percent of the eligible sample of 2225, or 1784 adults, agreed to be interviewed.

Variables

Information concerning each participant's education, occupation, and other health and psychosocial characteristics was obtained from a personal interview conducted in the home. Physical measurements included blood pressure, height, weight, and the circumference of the waist and hips. SES was defined by a composite index based on the participant's educational level and most recent occupation. A composite index was created because previous research in a neighboring county indicated that education alone was not an adequate measure of overall economic resources.[29] The number of years of formal education completed was divided into three categories: less than high school graduate, high school graduate, and greater than high school. Using Hollingshead rankings of job prestige,[30] occupation was grouped into unskilled, semiskilled, and skilled/white collar categories. Three levels of SES were created from the cross-classification of education and occupation categories (Table 1).

Dietary information was collected via an interviewer-administered food frequency questionnaire adapted from a standardized instrument developed by Block et al.[31] Dietary assessment focused on the portion size and the frequency with which certain foods were consumed during the previous year. The food list, portion size distributions, and nutrient composition values were derived from Block's standardized instrument.[31] The instrument was modified somewhat to reflect regional and ethnic food choices known to characterize southern Black populations.

In reliability studies, the correlation between repeated administrations of Block's standardized instrument[31] was 0.70 for each nutrient.[32,33] The validity of the standardized instrument was also judged to be adequate with nutrient correlations ranging from 0.5 to 0.7 when food frequency dietary data were compared to three 4-day food records.[33,34]

Food frequency data were converted to nutrients consumed per day (in mg) for each individual.[31] In addition to calculated values of energy, potassium, and calcium, two estimates of sodium intake were used. The first measure included only the sodium consumed in food. The second measure, enhanced sodium, included both food sodium and an estimate of discretionary table salt use. Table salt use was estimated as 0%, 15%, or 30%, respectively, of the individual's food sodium if they said they either rarely/never, sometimes, or often/always added salt to their food. Sodium ratios were also calculated. Individuals with missing data (n = 104) were excluded from the analysis.

Questions concerning other factors that may influence food choices and attitudes that reflect knowledge of diet and health principles included the number of times in a typical week that the participant ate out, the change in use of salt during the last 5 to 10 years (more, same, less), and perceptions about whether diet affects risk for disease.

Statistical Analysis

The relationship between nutrient intake and SES was assessed by estimating nutrient intake means within each socioeconomic status category by weighted covariate analysis[35] using a multiple linear regression model. Two indicator variables were used to categorize high and medium SES with low SES as the referent group. Men and women were analyzed separately because of the expected greater caloric intake among men. Weighted analyses were performed to take into account the sampling strategy as well as nonresponse. The multivariable analyses of the relation between SES and electrolytes also included age and total caloric intake.[36] Levels of other determinants of consumption and nutritional attitudes were also characterized by strata of socioeconomic status.

Results

Little difference was noted in mean age, waist-to-hip ratio (WHR), and energy intake by level of SES among men (Table 2). Body mass index (BMI) was higher among men with high SES than low SES. Lower SES women were older than those in the other groups. BMI, WHR, and energy intake decreased with increasing SES among women, possibly because of the differences in age among the groups. As expected, the mean energy intakes of men exceeded those of women.

Table 3 shows that sample mean intakes of sodium and enhanced sodium for men and women were near or above the upper limit of the "estimated safe and ad-

TABLE 2—Means and Standard Errors of Selected Characteristics by Sex and Socioeconomic Status (SES): Pitt County, NC, 1988

Characteristics	Total	SES1	SES2	SES3
Men				
Number	641	230	207	204
Weighted %		34.5	34.9	30.6
Age	35.0 ± 0.4	35.5 ± 0.6	34.7 ± 0.5	35.0 ± 0.7
BMI (kg/m^2)	26.3 ± 0.2	26.2 ± 0.3	26.1 ± 0.3	26.8 ± 0.3
WHR	0.89 ± 0.01	0.89 ± 0.01	0.89 ± 0.01	0.89 ± 0.01
Energy (kcal)	2427 ± 74	2451 ± 108	2530 ± 131	2283 ± 94
Women				
Number	1039	442	296	301
Weighted %		39.7	29.4	30.9
Age	35.5 ± 0.3	37.1 ± 0.5	34.6 ± 0.4	34.1 ± 0.4
BMI (kg/m^2)	29.4 ± 0.3	30.3 ± 0.5	29.1 ± 0.6	28.5 ± 0.4
WHR	0.85 ± 0.01	0.87 ± 0.01	0.85 ± 0.01	0.83 ± 0.01
Energy (kcal)	1822 ± 43	1918 ± 58	1784 ± 60	1736 ± 69

Note. SES1 = low; SES2 = middle; SES3 = high; BMI = body mass index; WHR = waist-to-hip ratio.

TABLE 3—Means and Standard Errors for Sodium, Potassium, and Calcium Intake by Sex, Comparing Pitt County, NC, Study Population with 1980 US RDAs for Adults[37]

Nutrients	Pitt County, NC		1980 RDAs
	Men	Women	
Sodium	3618 ± 116	2801 ± 67	1100–3300[a]
Enhanced sodium[b]	4310 ± 139	3175 ± 80	—
Potassium	2704 ± 79	2097 ± 45	1875–5625[a]
Calcium	664 ± 22	563 ± 14	800

Note. Intake is expressed as mg/day.
[a]Range of estimated safe and adequate daily intakes.
[b]Food sodium plus estimate of table salt use.

TABLE 4—Means of Sodium, Potassium, Calcium, Sodium to Potassium, and Sodium to Calcium Ratios, Adjusted for Age and Energy by Sex and SES: Pitt County, NC, 1988

Nutrients	SES1	SES2	P^a value	SES3	P^a value
Men					
Sodium	3694	3766	.20	3734	.34
Enhanced sodium[b]	4476	4469	.93	4377	.19
Potassium	2662	2794	.02	2877	< .001
Calcium	652	692	.05	709	.02
Sodium/potassium	1.42	1.36	.13	1.32	.003
Sodium/calcium	6.00	5.72	.23	5.65	.06
Women					
Sodium	2836	2847	.81	2861	.61
Enhanced sodium[b]	3252	3219	.48	3194	.38
Potassium	2062	2151	.02	2191	.004
Calcium	533	597	.005	597	< .001
Sodium/potassium	1.42	1.37	.11	1.35	.03
Sodium/calcium	5.67	5.45	.004	5.25	< .001

Note. SES1 = low; SES2 = middle; SES3 = high. Intake is expressed as mg/day.
[a]From one-degree of freedom *F*-statistic of beta coefficient from dummy variable for SES where SES1 is the referent.
[b]Food sodium plus estimate of table salt use.

equate daily dietary intakes" levels recommended as part of the 1980 Recommended Dietary Allowances (RDAs).[37] In addition, mean potassium was in the low range of recommended intake, particularly when considering the high general sodium intake and the fact that mean calcium intake fell below the RDA.

After adjusting for differences in age and energy intake among the SES groups, there was no consistent association between SES and sodium intake as measured by either food sodium or enhanced sodium (Table 4). Potassium and calcium intake increased with increasing SES for both men and women, with significantly ($P < .05$) higher intakes in the middle and high SES groups compared with the low SES group. Sodium to potassium and sodium to calcium ratios also decreased with increasing SES for both sexes. Similar results were found for the ratios when using enhanced sodium as the measure of sodium intake.

The distribution of potential determinants of nutrient consumption by level of SES are summarized in Table 5. Low SES respondents were more likely to live in large households (>5 persons). As expected, the proportion of men and women eating meals outside the home increased with increasing SES. For both sexes, those in the higher SES groups were more likely to agree with the statement that diet affects disease risk.

Men and women in the low SES group were more likely to report a high level of table salt use than those with higher SES. Men were more likely than women to salt food at the table. Both men and women in the low SES group were less likely to report a decrease in sodium intake from the previous 5 to 10 years. Interestingly, low SES men were less likely than middle and high SES men to perceive their salt intake as high. This last finding was not observed for women.

Discussion

Based on the 1980 RDA recommendations,[37] the distribution of electrolytes among Black adults in the Pitt County Study indicates lower than desirable calcium and potassium intakes and higher than desirable sodium intakes. Male study participants had a higher mean sodium intake and lower calcium and potassium intakes than 25- to 54-year-old Black and White men in NHANES II, which assessed dietary intake by 24-hour recall.[24] Sodium intake was also higher among women participants; however, calcium

and potassium intakes were similar. Sodium and potassium intakes in the Pitt County Study population were greater than those reported for Black men and women in Evans County in 1961.[38] A variety of factors may account for differences among these population studies, including secular trends, variation in dietary assessment instruments, and regional differences in food consumption.

Within the Pitt County Study population, Black men and women with high SES consumed more potassium and calcium and had lower sodium to potassium and sodium to calcium ratios than did those with low SES. These SES differences in electrolyte intake are consistent with the well-known inverse association between SES and hypertension.[4] However, the small magnitude of the differences in dietary intake across the SES categories (5% to 10%) in this study population is also consistent with literature reporting limited variation in dietary intake within populations compared to differences observed between populations.[16] Potential problems in the assessment of dietary intake include inflated variance estimates due to greater intra- vs interindividual variability. This could lead to an underestimation of the statistical significance of the differences among SES groups.[39]

Reasons that dietary intakes vary by SES are complex and involve many social and cultural processes. Not only do income, purchasing power, and access to food differ across SES, but access to and use of nutritional information probably differs by educational level. In addition, culture exerts a major influence on the selection, consumption, and use of the available food supply.[19] In this study, persons with high SES were more likely to believe that diet affects disease, to report less sodium consumption than in the past and less use of salt at the table. Approximately 70% of the study population did not consider their salt intake to be high. Given the overall high sodium consumption in this population, such a perception may warrant special attention in any community-based nutrition education program.

Large household size and the consumption of calories away from home have also been related to the intake of less healthy diets.[40,41] Compared to those with high SES, individuals with low SES were more likely to live in larger households but were less likely to consume foods away from home. The contrasting relations that household size (inverse) and the number of times respondents eat out per week (direct) have with SES may help explain the lack of variation in sodium intake across the groups.

TABLE 5—Percentage of Participants with Characteristics Affecting Food Consumption, by Sex and Socioeconomic Status (SES): Pitt County, NC, 1988

Characteristics	SES1	SES2	SES3
Men			
Household size (% >5)	20.2	11.4	8.7
Number of times eat out per week			
0	29.3	18.5	12.8
1	29.7	22.5	26.2
>1	41.0	58.3	60.0
Believe diet affects disease risk	77.3	91.2	91.8
Use tablet salt often/always	35.6	28.6	27.3
Consume less salt compared to 5–10 years ago	30.7	49.6	54.6
Perceive salt intake as high	23.1	28.8	32.0
Women			
Household size (% >5)	23.0	13.3	8.8
Number of times eat out per week			
0	40.3	23.3	14.2
1	33.2	34.4	22.5
>1	26.4	42.3	63.3
Believe diet affects disease risk	82.3	89.7	93.4
Use table salt often/always	23.5	21.0	19.2
Consume less salt compared to 5–10 years ago	45.2	52.7	52.2
Perceive salt intake as high	30.9	32.0	32.9

Note. SES1 = low; SES2 = middle; SES3 = high.

In conclusion, the differences in electrolyte intake by socioeconomic status observed in this population are consistent with the frequently reported inverse association between SES and hypertension. In this study population, a large percentage of individuals believed that diet influences risk for disease, but a much smaller percentage consumed the recommended level of these nutrients, reported reducing their salt intake over time, or considered their current salt intake to be too high. Therefore, health promotion efforts that target dietary behaviors in this and similar communities may need to develop population-based strategies to change food consumption patterns in the community as a whole. This will require an understanding of the social and cultural norms that sustain certain dietary behaviors at the individual level as well as an understanding of how local economic factors influence dietary options. Public health strategies that more effectively educate high-risk populations about the role of diet in cardiovascular diseases are clearly warranted. □

Acknowledgments

This research is supported by grant HL33211 from the National Heart, Lung, and Blood Institute.

This paper was presented in part at the 5th International Interdisciplinary Conference on Hypertension in Blacks, May 3–7, 1990, Long Beach, Calif.

The authors thank Janet B. Croft, Brian Rice, Steven R. Browning, and Bob Kleckner for their valuable contributions to this work.

References

1. Prineas RJ, Gillum RF. U.S. epidemiology of hypertension in Blacks. In: Hall WD, Saunders E, Shulman NB, eds. *Hypertension in Blacks: Epidemiology, Pathophysiology, and Treatment.* Chicago, Ill: Yearbook Medical Publishers; 1985:17–36.
2. James SA. Psychosocial and environmental factors in Black hypertension. In: Hall WD, Saunders E, Shulman NB, eds. *Hypertension in Blacks: Epidemiology, Pathophysiology, and Treatment.* Chicago, Ill: Yearbook Medical Publishers; 1985:132–143.
3. James SA. Psychosocial precursors of hypertension: a review of the epidemiologic evidence. *Circulation.* 1987;76(suppl):160–165.
4. Hypertension Detection and Follow-up Program (HDFP) Cooperative Group. Race, education, and prevalence of hypertension. *Am J Epidemiol.* 1977;106:351–361.
5. Harburg E, Erfurt JC, Chape C, Hauenstein LS, Schull WJ, Schork MA. Socio-ecological stressor areas and Black-White blood pressure: Detroit. *J Chronic Dis.* 1973;26:595–611.
6. Dischinger PC, Apostolides AY, Entwisle G, Hebel JT. Hypertension incidence in a inner-city Black population. *J Chronic Dis.* 1981;34:405–413.
7. Syme SL, Oakes TW, Friedman GD, Feldman R, Siegelaub AB, Collen M. Social

class and racial differences in blood pressure. *Am J Public Health.* 1974;64:619–620.
8. Keil JE, Tyroler HA, Sandifer SH, Boyle EJ. Hypertension: effects of social class and racial admixture: the results of a cohort study in the Black population of Charleston, SC. *Am J Public Health.* 1977;67:634–639.
9. James SA, LaCroix AZ, Kleinbaum DG, Strogatz DS. John Henryism and blood pressure differences among Black men, II: the role of occupational stressors. *J Behav Med.* 1984;7:259–275.
10. Strogatz DS, James SA, Haines PS, et al. Alcohol consumption and blood pressure in Black adults: the Pitt County Study. *Am J Epidemiol.* 1991;133:442–450.
11. Mason TJ, Fraumeni JR, Hoover R, Blot WJ. An Atlas of Mortality from Selected Diseases. Washington, DC: Government Printing Office; 1981. US Dept of Health and Human Services publication (NIH) 81-2397.
12. Horan M, Blanstein W, Dunbar J, et al., eds. *NIH Workshop on Nutrition and Hypertension Proceedings.* New York, NY. Biomedical Information Corporation; 1985.
13. Harlan WR, Hull AL, Schmouder RL, Landis JR, Thompson FE, Larkin FA. Blood pressure and nutrition in adults: the National Health and Nutrition Examination Survey. *Am J Epidemiol.* 1984;120:17–28.
14. Gruchow H, Sobocinski KA, Barboriak J. Alcohol, nutrient intake, and hypertension in U.S. adults. *JAMA.* 1985;253:1567–1570.
15. McCarron D, Morris C, Henry H, Stanton JL. Blood pressure and nutrient intake in the United States. *Science.* 1984;224:1392–1397.
16. Freis EP. Salt volume and the prevention of hypertension. *Circulation.* 1976;53:589–595.
17. Dawber TR, Kannel WB, Kagan A, Donalbedian RK, McNamara P, Pearson G. Environmental factors in hypertension. In: Stamler J, Stamler R, Pullman T, eds. *The Epidemiology of Hypertension.* New York, NY: Grune and Stratton; 1967:255–288.
18. INTERSALT Cooperative Research Group. INTERSALT: an international study of electrolyte excretion and blood pressure: results for 24-hour urinary sodium and potassium excretion. *Br Med J.* 1988;297:319–328.
19. Axelson ML. The impact of culture on food-related behavior. *Ann Rev Nutr.* 1986;6:345–363.
20. Popkin BM, Haines PS. Factors affecting food selection: the role of economics. *J Am Diet Assoc.* 1981;79:415–425.
21. Kushi LH, Folsom AR, Jacobs DR, Luepker RV, Elmer PJ, Blackburn H. Educational attainment and nutrient composition patterns: the Minnesota Heart Survey. *J Am Diet Assoc.* 1988;58:1230–1236.
22. Windham CT, Wyse BW, Hanson G, Hurst RL. Nutrient density of diets in the USDA Nationwide Food Consumption Survey, 1977–1978, I: impact of socioeconomic status on dietary density. *J Am Diet Assoc.* 1983;82:28–34.
23. Cronin FJ, Krebs-Smith SM, Wyse BW, Light L. Characterizing food usage by demographic variables. *J Am Diet Assoc.* 1982;81:661–672.
24. National Center for Health Statistics, Carrol MD, Abram S, Dresser CM. *Dietary intake source data: United States, 1976–1980. From Vital and Health Statistics* series 11, No. 231, Marc., 1983. DHHS Pub No (PHS)83-1681. Public Health Service. Washington, DC. US Government Printing Office.
25. National Center for Health Statistics. Selected findings: food consumption profiles of White and Black persons 1–74 years of age in the U.S., 1971–74. *Advanced Data.* Pub. 21. June 1978.
26. USDA Nationwide Food Consumption Survey. *Food Intakes: Individuals in 48 States, 1977–78.* Hyattsville, Md: Consumer Nutrition Division, Human Nutrition Information Service; 1983. Report No.I-1.
27. Goodhart RS, Shils M. *Modern Nutrition in Health and Disease.* 6th ed. Philadelphia, Pa: Lea and Febiger; 1980.
28. Langford HG, Langford FPJ, Tyler M. Dietary profile of sodium, potassium, and calcium in U.S. Blacks. In: Hall WD, Saunders E, Shulman NB, eds. *Hypertension in Blacks: Epidemiology, Pathophysiology, and Treatment.* Chicago, Ill: Yearbook Medical Publishers; 1985:49–57.
29. James SA, Strogatz DS, Wing SB, Ramsey DL. Socioeconomic status, John Henryism and hypertension in blacks and whites. *Am J Epidemiol.* 1987;126:664–673.
30. Hollingshead AB. Four factor index of social status (unpublished working paper). New Haven, Conn: Department of Sociology, Yale University; 1975.
31. Block G, Hartman AM, Dresser CM, Carroll MD, Gannon J, Gardner L. A data based approach to diet questionnaire design and testing. *Am J Epidemiol.* 1986;24:452–469.
32. Leighton J, Neuget AI, Block G. A comparison of face-to-face food frequency interviews with telephone interviews and self-administered questionnaires. *Am J Epidemiol.* 1988;128:89. Abstract.
33. Block G, Hartman A. Issues in reproducibility and validity of dietary studies. *Am J Clin Nutr.* 1989;50:1133–1138.
34. Block G, Hartman A, Naughton D. A reduced dietary questionnaire: development and validation. *Epidemiology.* 1990;1:58–64.
35. Kleinbaum DG, Kupper LL, Muller KE. *Applied Regression Analysis and Other Multivariable Methods.* Boston, Mass: *PWS-Kent Pub. Com.;* 1987.
36. Willet WC, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. *Am J Epidemiol.* 1986;124:17–27.
37. Committee on Dietary Allowances, National Research Council. *Recommended Dietary Allowances.* 9th rev ed. Washington, DC: National Academy of Sciences; 1980.
38. Grim CE, Luft FC, Miller JZ, et al. Racial differences in blood pressure in Evans County, Georgia: relationship to sodium and potassium intake and plasma renin activity. *J Chronic Dis.* 1980;133:87–94.
39. Liu K, Stamler J, Dyer A, McKeever J, McKeever P. Statistical methods to assess and minimize the role of intra-individual variability in observing the relationship between dietary lipids and serum cholesterol. *J Chronic Dis.* 1978;31:399–419.
40. Harris EW, Randall E. The influence of head of household structure on the dietary quality of food available for consumption inside U.S. households, 1977–1978. *Ecol Food N.* 1987;20:109–120.
41. Bunch K, Hall L. *Factors Affecting Nutrient Consumption.* Ithaca, NY: Cornell University; 1983. A.E.Res.83–32.

Representativeness of Deaths Identified through the Injury-at-Work Item on the Death Certificate: Implications for Surveillance

Julie Russell, PhD, MPH, and Carol Conroy, PhD, MPH

ABSTRACT

Background. This research investigated the accuracy of the injury-at-work item on the death certificate for surveillance of occupational injury deaths in Oklahoma during 1985 and 1986.

Methods. Representativeness of occupational injury deaths identified by death certificates was assessed by comparing these deaths with all occupational injury deaths identified through death certificates, workers' compensation reports, medical examiner reports, and OSHA records for categories of occupation, industry, and external causes of death.

Results. Certain external causes of death (e.g., motor vehicle traffic deaths) and certain occupations (e.g., farming) and industries (agriculture and services) are more often underidentified through death certificates.

Conclusions. The findings of this study support Baker's observation that no single data source contains all deaths or all the data elements necessary to describe occupational injury deaths. Data sources may be combined to improve representativeness through more complete case ascertainment. (*Am J Public Health.* 1991;81:1613–1618)

Introduction

Injury is the leading cause of death for persons under the age of 45, accounting for more years of potential life lost than heart disease and cancer combined.[1] Occupational injuries are an important subset of the larger injury problem and are estimated to account for one sixth of all fatal injuries to persons between the ages of 17 and 64.[2] Several private and federal agencies independently estimate the number of occupational injury deaths each year; however, there is currently no agreed-upon method for making such estimates.[3–7] For this reason, the annual number of occupational injury deaths differs widely between reporting sources. For example, the following widely discrepant estimates were made for the number of occupational injury deaths occurring in 1985: Bureau of Labor Statistics, 3750; National Institute for Occupational Safety and Health (NIOSH), 6385; and National Safety Council, 11 170.[3,5,6]

Several states and countries have used death certificates to describe occupational injury deaths.[2,3,8,9–28] They have also been used alone and in conjunction with other sources to study fatal injuries to farmers[22–24] and truck drivers[29] and homicides and suicides in the workplace.[14,25,26] The Division of Safety Research at NIOSH uses death certificates for surveillance of occupational injury deaths.[3]

The only variable on the death certificate used by NIOSH to define work relatedness is the injury-at-work item. However, the National Center for Health Statistics (NCHS) does not include this item in its vital statistics mortality tapes, there is no national guideline for its completion, and different states may define injury at work differently.

Table 1 summarizes the results of several state studies comparing case ascertainment using various data sources. These studies demonstrate that death certificates identify between 57% and 88% of occupational injury deaths. The literature indicates that occupational injury deaths are underreported using death certificates and that certain industries, occupations, external causes of death, and gender or race groups may consistently be underrepresented. For example, deaths that occur in occupations involving substantial amounts of driving of motor vehicles on public roads are often not identified as occupational fatalities.[8,29–31]

While surveillance systems often rely on sources that do not identify every occurrence of the condition under study,[32–35] the cases that are captured must be representative of the population of cases. If missing cases are distributed in occupation and industry groups similarly to cases in the surveillance system, there may be no bias in conclusions based on these data; however, if the cases missed are systematically different from those identified, then their exclusion could bias results, causing public health resources to be misdirected. Therefore, the representativeness of deaths identified by review of the injury-at-work item on death certificates is

At the time of this study, the authors were with the National Institute for Occupational Safety and Health, Division of Safety Research, Morgantown, WVa. Julie Russell is now with the Division of Injury Control, Centers for Disease Control, Atlanta, and Carol Conroy is now with the California Occupational Health Program, Berkeley.

Requests for reprints should be sent to Julie Russell, PhD, DIC-CDC-F-36, 1600 Clifton Road, Atlanta, GA 30033.

This paper was submitted to the journal December 7, 1990, and accepted with revisions May 28, 1991.

TABLE 1—Percentage of Work-Related Fatal Injuries Identified from Various Data Sources, Selected Studies, United States, 1976–1987

Author, Date	State, Year(s)	Data Sources	Death Certificates Only	
			All Deaths (%)	Motor Vehicle Deaths (%)
Karlson, 1978[29]	Wisconsin, 1976	DC, WC	76	60
Baker,[a] 1982[31]	Maryland, 1978	DC, WC, ME, OSHA	67	
Roberts, 1983[9]	Michigan, 1980	DC, SIC	79	56
Paulson, 1986[30]	Colorado, 1982–1984	DC, WC, OSHA, MSHA	86	67
Robinson,[b] 1988[27]	Allegheny County, Pa, 1979–1982	DC, ME	57	
Friedman, 1989[28]	Massachusetts, 1987	DC, WC, OSHA, SIC	88	73

Note. DC = death certificates; ME = medical examiner reports; WC = workers' compensation reports, OSHA = Occupational Safety and Health Administration fatality/catastrophe reports; SIC = State Industrial Commission reports; MSHA = Mine Safety and Health Administration.
[a]Baker's study did not report the exact number of motor vehicle deaths.
[b]Robinson's study excluded motor vehicle injury deaths.

an important issue. This study was undertaken to evaluate the representativeness of deaths identified through death certificates alone for one vital statistics reporting unit (Oklahoma).

Methods

The study population consisted of all (resident and nonresident) workers who were fatally injured while working in Oklahoma during calendar years 1985 and 1986. Most of the analysis is restricted to males because injuries to females were too few to statistically analyze, although a brief description of these deaths is included in the results.

Work-related injury deaths were identified from four primary data sources: death certificates, Occupational Safety and Health Administration (OSHA) fatality/catastrophe reports, workers' compensation reports, and medical examiner reports. Two secondary data sources, workers' compensation court records and police accident reports, were used to confirm work relatedness when it could not be determined from the primary data sources. The working population was estimated from the Bureau of Labor Statistics Geographic Profiles of Employment and Unemployment for 1985 and 1986.[36,37]

An injury death was defined as an intentional or unintentional death resulting from an injury that can be classified by International Classification of Diseases (9th revision; ICD-9) external cause of death codes ranging from E800 to E999.[38] Work related was defined as including full-time or part-time workers as well as unpaid family members who are engaged in work activities in a work environment (e.g., family farms or other family businesses) at the time of their death.[2,39]

Death certificates on which the injury-at-work item was marked "yes" were acquired from the State Vital Statistics Office, as well as death certificates for work-related deaths that were identified from other sources and were not marked "yes" for injury at work.

Copies of the OSHA fatality/catastrophe reports for 1985 and 1986 were obtained from the state OSHA office. These reports are completed by OSHA within 48 hours of a death, the time in which employers are required to report a death to OSHA. Those regulated by other federal safety and health laws, government employers, and domestic workers are excluded from reporting requirements.

A listing for all filed workers' compensation death claims for 1985 and 1986 in Oklahoma was obtained from the State Workers' Compensation Court. These claims do not include self-employed individuals and government and domestic workers who are not eligible for workers' compensation benefits.

In most states, as in Oklahoma, medical examiners or coroners have jurisdiction over all occupational injury deaths.[40] Hand searching of medical examiners' records was required to identify occupational injury deaths in Oklahoma. Records of deaths were then matched visually by agreement of location of death, sex, external cause of death, date of birth, and name.

Demographic and external cause of death data were abstracted from death certificates. Results for the death certificate (DC) group were classified by "usual" occupation/industry, since only usual occupation/industry is reported on the death certificate. This represents the job that was worked for the longest period of time. The comparison group, all deaths (AD), was classified by the most current occupation/industry information that could be derived from any of the data sources. When workers' compensation and OSHA records were available, they were used as the primary sources of current occupation/industry information. When these reports were not available, the medical examiner reports and death certificates were studied in detail to determine whether the "usual" occupation/industry reported on the death certificate was the same as the current occupation/industry. The occupation/industry reported on the death certificate was determined to not be current when it was not consistent with the location and external cause of death. For example, a death certificate may indicate an occupation/industry of student/school, while the injury description on the death certificate and the medical examiner report indicate that the death occurred from a fall at a grain elevator while the victim was working. In this case, there is clear evidence that the "usual" occupation/industry information does not reflect the current occupation/industry.

Annual average rates were calculated by dividing the number of deaths for an occupation or industry group by the working population (i.e., the number of people employed in 1985 and 1986 in each occupation or industry group). Rate ratios were used to compare the pooled data rate and the death certificate–based rate. Rate ratios were calculated for occupation and industry for AD and DC.

Results

Three hundred twenty-nine occupational injury deaths (299 males, 30 females) were identified during 1985 and 1986, making the annual average death rate 11.3 per 100 000 workers (2.3 for females, 18.7 for males). Twenty-eight of the 30 (93%) female occupational injury deaths were identified by death certificates. Seventy percent of the female deaths occurred in two workplace disasters: an explosion in a fireworks factory and a mass homicide at a post office. These events were highly publicized, possibly increasing the chances of the deaths being recorded as work related. Seventy-three percent of the female deaths were in the manufacturing and public administration industries. Sixty-seven percent of female deaths were in the laborers and helpers and administrative support occupations. There were no female deaths in the four highest risk industries of transportation, communications, and public utilities; construction; agriculture, forestry, and fishing; and mining. Females were excluded from further analysis because of the small numbers of deaths and the two clusters of female deaths.

Of the 299 male deaths, 82% (244) were identified through medical examiner reports, 72% (215) through death certificates, 57% (169) through workers' compensation reports, and 24% (72) through OSHA fatality/catastrophe reports (Figure 1). Cases identified through OSHA fatality/catastrophe reports were not included in Figure 1 because only one additional case was reported to OSHA and not identified in the other three data sources.

There was 86% agreement between usual and current industry for the two-digit Standard Industrial Classification (SIC) and 81% agreement between usual and current occupation when 11 broad occupation categories were used.[41]

When characteristics of deaths were compared for AD and DC, the distributions for age, race, and place of injury were generally similar. The distributions for external cause, shown in Table 2, are similar except for a discrepancy in the identification of traffic deaths (ICD-9 E810.0 to E819.9) between AD and DC. Only 60% (44/73) of these deaths were identified through death certificates (Figure 2).

The distribution of intentional injuries was similar between AD and DC except for suicides that occurred at work, with only 14% (2/14) of suicides being identified through death certificates. Since suicides account for only 5% of AD, the proportion in relation to all occupational injury deaths did not change.

Rate ratios are used to compare industry death rates for AD and DC in Table 3. Rate ratios are 1.5 or greater for four industries: finance, insurance, and real estate; agriculture, forestry, and fishing;

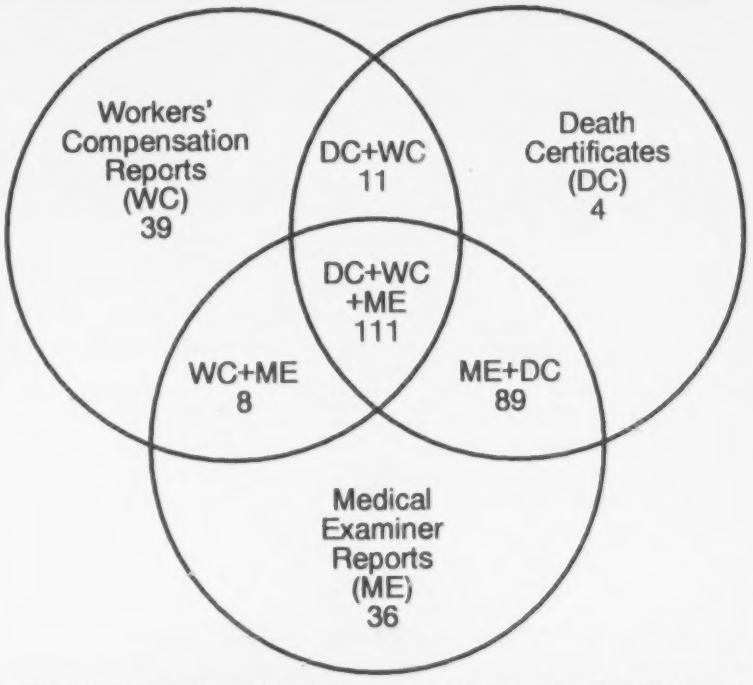


FIGURE 1—Number of cases identified, by each data source, male work-related deaths, Oklahoma, 1985–1986.

TABLE 2—External Causes of Death by Ascertainment Source: Male Work-Related Deaths, Oklahoma, 1985–1986

ICD-9 E Code	DC (frequency)	AD[a] (frequency)	DC/AD (%)
Traffic (E810–E819)	44	73	60
Motor Vehicle Non-Traffic (E820–E825)	8	8	100
Air and Other Transportation (E826–E848)	8	11	73
Falls (E880–E888)	12	14	81
Fires (E890–E899)	7	8	88
Environment, Submersion, Suffocation (E900–E915)	9	13	69
Blunt Trauma (E916–E918)	14	15	93
Electrocution (E925)	26	30	87
Explosive Material (E923)	11	12	92
Machines (E919)	35	43	81
Other[b]	11	24	46
Suicide (E950–E959)	2	14	14
Homicide (E960–E969)	28	31	90
Total	215	296	

Note. ICD-9 = *International Classification of Diseases*, 9th revision; DC = death certificates; AD = all deaths.

[a]Includes cases identified by death certificates, medical examiner reports, Occupational Safety and Health Administration fatality/catastrophe reports, and workers' compensation reports.

[b]Other includes Poisoning (E850–E869), Explosions (E921), Legal Intervention (E970–E978), Late Effects of Injury (E929), and Other Accidental Causes (E928).

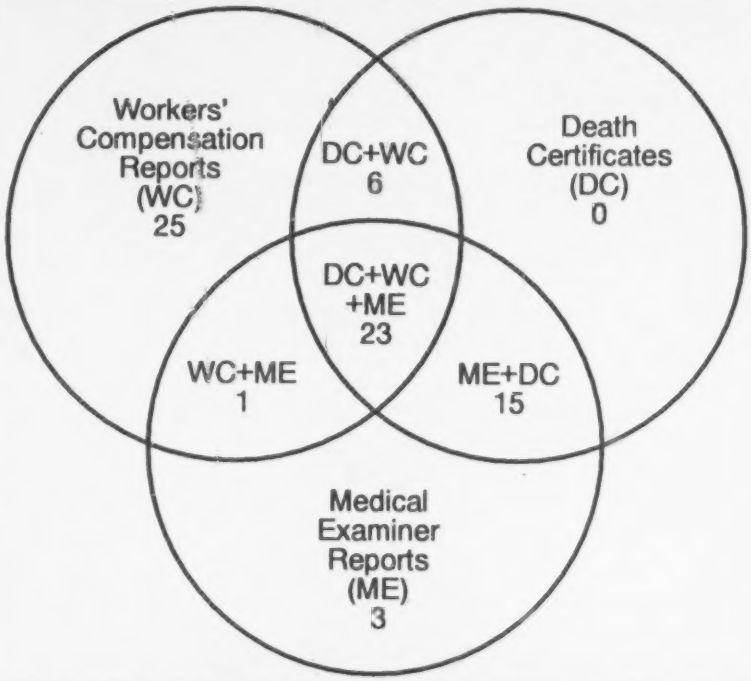


FIGURE 2—Number of cases identified, by each data source, traffic deaths, male work-related deaths, Oklahoma, 1985—1986.

TABLE 3—Annual Average Death Rates per 100 000 Workers by Industry Groups and Ascertainment Source: Male Work-Related Deaths, Oklahoma, 1985–1986

Industry Group	Usual Industry Death Certificate Rate	Current Industry All Death[a] Rate	Rate Ratio
Transportation, Communication, Utilities (SIC 400–499)	46.4	52.3	1.1
Mining (SIC 100–149)	37.5	56.8	1.5
Construction (SIC 150–179)	30.4	33.2	1.1
Agriculture, Forestry, Fishing (SIC 0–99)	14.9	34.0	2.3
Manufacturing (SIC 200–399)	6.5	9.4	1.4
Public Administration (SIC 910–979)	9.2	11.0	1.2
Services (SIC 700–899)	6.4	14.7	2.3
Trade (SIC 500–599)	6.0	11.3	1.9
Finance, Insurance, Real Estate (SIC 600–679)	2.0	7.8	3.9
Overall rate	15.4	21.4	1.4

Note. Industry groups are defined by the Office of Management and Budget, Standard Industrial Classification (SIC), 1987.
[a]Includes cases identified by death certificates, medical examiner reports, Occupational Safety and Health Administration fatality/catastrophe reports, and workers' compensation reports.

services; and mining. Fifty-one percent of the deaths not identified through death certificates occurred in three industries (mining, agriculture, and services).

Occupation death rates are compared for AD and DC in Table 4. Rate ratios were 1.5 or greater for five occupation categories: helpers and laborers; farming occupations; executive, administrative, and managerial occupations; technical and related support; and machine operators, assemblers, and inspectors. Fifty-seven percent of the deaths missed by death certificates were in the four highest risk occupations (helpers and laborers; transportation and material moving operatives; farming occupations; and precision production, crafts, and repairers).

In Figure 3 the distributions for death rates by industry based on the pooled data are compared with the death rates based on death certificates only. The distributions have the same shape; however, the pooled data show higher death rates per 100 000 workers and substantially higher rates for mining and agriculture.

Discussion

Death certificates are shown in this research to identify 72% of the fatal occupational injuries occurring to males and 93% for females in Oklahoma during 1985 and 1986. Previous studies show that death certificates alone underestimate the magnitude of the occupational injury problem in the United States by 24% to 33%.[2,30] Although the results of this study (underestimation by 28% for males) cannot be directly generalized to the United States, they agree with results of previous studies in Colorado,[30] Maryland,[2] and Wisconsin[29] on the basis of case ascertainment using death certificates.

Certain external causes of injuries (traffic deaths) and certain industries (mining, agriculture, and services) and occupations (laborers and helpers; farmers; transportation and material moving operatives; and precision production, crafts, and repairers) were less often identified through the injury-at-work item on the death certificate than others and point to areas where estimates of death rates based on death certificates should be interpreted with caution.

Death certificates have been shown to accurately identify broad categories of occupation and industry that are at highest risk of occupational injury death when NIOSH's death certificate database was compared with the Bureau of Labor Statistics annual survey.[19] While this broad agreement is useful to draw attention to the problem and to estimate its overall magnitude, more targeted surveillance for specific external causes or occupation/industry groups is necessary to plan and implement intervention strategies. This targeted surveillance may not be possible through death certificates.

The use of death certificate data for surveillance of occupational injury deaths hinges on the assumption that usual occupation and industry, as reported on the

death certificate, serves as an acceptable proxy for the occupation/industry the person was engaged in at the time of death (current occupation and industry). This assumption appears to be valid for analysis of broad categories of industry and occupation. However, lack of current occupation and industry information is a serious limitation to any study based on death certificate data alone and intended to target interventions in a defined population (e.g., a county or state), since the occupational and industry groups at highest risk are the least likely to be identified through death certificates.

Occupation/industry is a crude method to measure the exposure of workers to harmful energy exchanges at work; therefore, accurate information on the current occupation/industry of the decedent at the time of death is critical to planning and implementation of prevention programs. Two recent studies show that the agreement for usual industry and current industry ranges from 68% to 72% and the agreement for usual occupation and current occupation ranges from 70% to 78% for the identified occupational injury deaths.[30,42] Goodman[22] found that 59% of people killed on farms in Georgia during 1971 to 1981 did not have the occupation of farmer noted on their death certificate. Furthermore, a group of potentially high-risk workers—students—are often not classified by the industry or occupation they were working in at the time of death. National death certificate data from 1980 to 1985 reveal that 16% of the deaths of persons under the age of 21 were noted with industry as student on the death certificate. The proper classification of this group of workers by the occupation they were engaged in at the time of fatal injury is important to the surveillance and prevention of injury deaths to young workers.

National surveillance of occupational injury deaths is needed to provide national estimates of fatal occupational injuries by occupation, industry, and external cause of death. These estimates can help prioritize research programs and target specific occupations, industries, and causes of death for intervention programs. Because death certificates are the only source of data that potentially includes all workers, every effort should be made to improve identification of occupational injury deaths on death certificates and improve the quality of the occupation and industry data, in addition to developing systems to link data sources for national surveillance. The findings of this study support the observation that no single data source contains all of the data elements needed to describe occupational injury deaths and plan prevention strategies.[2,43]

Measures to improve representativeness as well as ascertainment of occupational deaths through the use of death certificates include the following:

1. A clear definition of when a person is considered "at work" is needed for recording work-related injuries that can be used in workers' compensation reports and by coroners, medical examiners, and researchers. There is also a need for a definition of a work-related injury. For example, a woman who is killed by her husband while she is working suffered an "injury at work"; however, many death certifiers may be reluctant to call the death work related.

2. Certain classes of ICD-9 external

TABLE 4—Annual Average Death Rates per 100 000 Workers by Occupational Group and Ascertainment Source: Male Work-Related Deaths, Oklahoma, 1985–1986

Occupational Group	Usual Occupation Death Certificate Rate	Current Occupation All Death[a] Rate	Rate Ratio
Helpers and Laborers (OCS 863–889)	36.0	57.2	1.6
Transportation and Material Moving (OCS 803–859)	36.6	46.4	1.3
Farming (OCS 473–499)	15.3	31.6	2.1
Precision Product, Craft, Repair (OCS 503–699)	11.7	15.4	1.3
Services (OCS 403–469)	10.0	13.7	1.4
Sales (OCS 243–285)	6.6	9.3	1.4
Executive, Administrative, Managerial (OCS 003–037)	3.3	7.6	2.3
Technical Support (OCS 203–235)	5.5	12.9	2.3
Professional Specialty (OCS 043–199)	5.4	7.8	1.4
Administrative Support (OCS 303–389)	11.7	10.6	1.1
Machine Operators, Assemblers, Inspectors (OCS 703–799)	3.0	6.0	2.0
Overall rate	13.5[b]	18.7[c]	1.4

Note. Occupational groups are defined by the 1980 Bureau of Census Occupation Coding System (OCS).
[a]Includes cases identified by death certificates, medical examiner reports, Occupational Safety and Health Administration fatality/catastrophe reports, and workers' compensation reports.
[b]Twenty-four cases identified by death certificates had occupational information missing or not classified.
[c]Fifteen cases in the pooled data had occupational information missing or not classified.

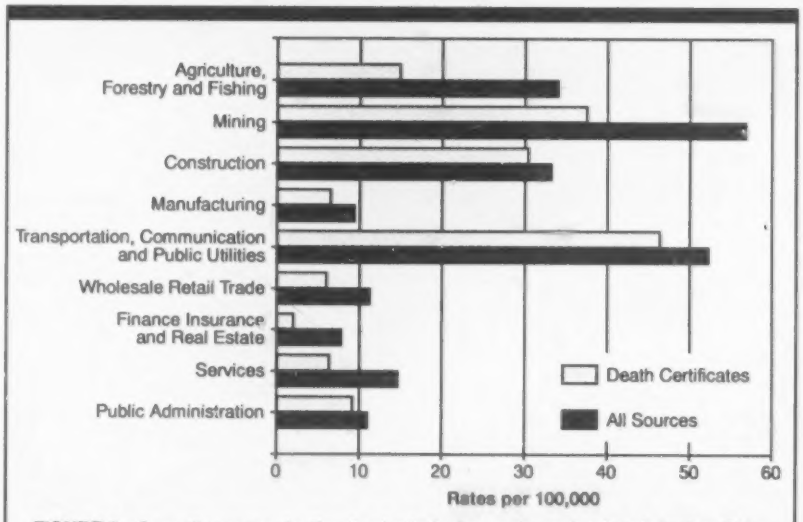


FIGURE 3—Annual average death rates by industry, male work-related deaths, Oklahoma, 1985–1986.

cause of death codes that are usually occupational (e.g., ICD-9 E919) should be identified, and state vital statistics offices should be asked to query death certifiers to confirm work relatedness when the injury at work item is marked "no" or is blank.

3. Current industry and occupation, in addition to "usual" occupation and industry, should be included on the US standard death certificate.

4. Responsibility for identifying and automating death certificates for injury-at-work deaths should be transferred to NCHS, so there will be uniform recording, quality control procedures, and timely dissemination of data.

5. Death certificates should be linked with medical examiner/coroner records in states that have a centralized and automated medical examiner/coroner system to make toxicology screening data available as well as detailed information on the nature and severity of injuries.

6. Death certificates should be linked with workers' compensation first reports of injury to improve case ascertainment, especially of occupational motor vehicle injury deaths. □

Acknowledgments

This research was supported by the National Institute for Occupational Safety and Health, Division of Safety Research.

Results of this research were presented, in part, at the 116th Annual Meeting of the American Public Health Association, in Boston, Mass., November 1988.

We are indebted to Drs. Greg Istre and Mark Roberts of the Oklahoma State Department of Health and to the Oklahoma chief medical examiner, Dr. Fred Jordan, and his staff. We also acknowledge the cooperation of the Oklahoma Workers' Compensation Court and the Oklahoma Occupational Safety and Health Administration, who provided records and helped make this research possible.

References

1. Committee on Trauma Research. *Injury in America: A Continuing Public Health Problem.* Washington, DC: National Academy Press; 1985.
2. Baker PS, Samkoff JS, Fisher RS, VanBuren CB. Fatal occupational injuries. *JAMA.* 1982;248:692–697.
3. Bell CA, Stout NA, Bender TA, et al. Fatal occupational injuries in the United States, 1980–1985. *JAMA.* 1990;263:3047–3050.
4. Pollack ES, Keimig DG, eds. *Counting Injuries and Illness in the Work Place: Proposals for a Better System.* Washington, DC: National Academy Press; 1987.
5. National Safety Council. *Accident Facts*; 1985.
6. Bureau of Labor Statistics. *Occupational Injuries and Illnesses in the US by Industry for 1985.* Washington, DC: Bureau of Labor Statistics; 1986.
7. Bureau of Labor Statistics. *Occupational Injuries and Illnesses in the US by Industry for 1986.* Washington, DC: Bureau of Labor Statistics; 1987.
8. Roberts C, Eyster J. *A Comparison of 1980 Fatality Reports and Death Certificates in Michigan.* Office of Vital and Health Statistics, Lansing, Mich.: Michigan Department of Public Health; 1983.
9. Pennsylvania Department of Health. *Mortality Experience of Pennsylvania Workers, 1983.* Harrisburg, Pa: Pennsylvania State Health Data Center; 1985.
10. Centers for Disease Control. Fatal occupational injuries in Texas, 1982. *MMWR.* 1985;34:130–134.
11. Honchar P, Suarez L. Death certificates as a surveillance tool for fatal occupational injuries. Presented at the V International Symposium on Epidemiology and Occupational Health, Sept. 9-11, 1986: Los Angeles, Calif. 1986.
12. Parkinson DK, Grauss WF, Perper JA, Elliot SA. Traumatic work place deaths in Allegheny County, Pennsylvania, 1983 and 1984. *J Occup Med.* 1986;28:100–102.
13. Davis H. Work place homicides of Texas males. *Am J Public Health.* 1987;77:1290–1293.
14. Davis H, Honchar PA, Suarez L. Fatal occupational injuries of women, Texas 1975–84. *Am J Public Health.* 1987;77:1524–1527.
15. Kizer K. *California Occupational Mortality 1979–1981.* Sacramento, Calif: California Department of Health Services; 1987.
16. Marine WM. Colorado Population-Based Occupational Injury and Fatality Surveillance System Report, 1982–1984. Presented at NIOSH Symposium, May 1987; Beckley, West Virginia.
17. Cone JE, Daponte A, Becker C, Reiter R, Balmes J. *Fatal Occupational Injuries in California.* San Francisco, Calif: California Department of Health Services; 1988.
18. Cryer PC, Fleming C. A review of work-related fatal injuries in New Zealand 1975–84—numbers, rates and trends. *NZ Med J.* 1987;100:1–6.
19. Stout-Wiegand N. Fatal occupational injuries in US industries, 1984. Comparison of two national surveillance systems. *Am J Public Health.* 1988;78:1215–1217.
20. Stout-Wiegand N. Fatal occupational injuries in the United States in 1980–1984. *Scand J Work Environ Health.* 1988;14:90–92.
21. Suruda A, Emmett EA. Counting recognized occupational deaths in the United States. *J Occup Med.* 1988;11:868–872.
22. Goodman RA, Smith JD, Sikes KR, Rogers DL, Mickey JL. Fatalities associated with farm tractor injuries: an epidemiologic study. *Public Health Rep.* 1985;100:329–333.
23. Helgerson SD, Milham S. Farm workers electrocuted when irrigation pipes contact powerlines. *Public Health Rep.* 1985; 100:325–328.
24. Murphy DJ, Seltzer BL, Yesalis CE. Comparison of two methodologies to measure agricultural occupational fatalities. *Am J Public Health.* 1990;80:198–199.
25. Kraus JF. Homicide while at work: persons, industries and occupations at high risk. *Am J Public Health.* 1987;77:1285–1289.
26. Conroy C. Suicide in the workplace: Incidence, victim characteristics, and external cause of death. *J Occup Med.* 1989;31:847–851.
27. Robinson CC, Kuller LH, Perper JA. An epidemiologic study of sudden death at work in an industrial county, 1979–1982. *Am J Epidemiol.* 1988;128:806–820.
28. Friedman DJ, Cohen BB, Davis L. *Traumatic Occupational Fatalities in Massachusetts, November 1986–October 1987.* Boston, Mass: Massachusetts Department of Public Health; 1989.
29. Karlson TA, Baker PS. Fatal occupational injuries associated with motor vehicles. *Am Assoc Automotive Med.* 1978;1:229–241.
30. Paulson KM, Marine WM, Weiser JK, Keefer SM. *Colorado Population-Based Occupational Injury and Fatality Surveillance System Report, 1982–1984.* Denver, Colo: Colorado Department of Health; 1988.
31. Baker PS, Wong J, Boron R. Professional drivers: protection needed for a high risk occupation. *Am J Public Health.* 1976;66:649–654.
32. Ing RT, Baker SP, Ellee JB, Frankowski RR, Guyer B, et al. Injury surveillance systems—strengths, weaknesses, and issues workshop. *Public Health Rep.* 1985;100:582–586.
33. Graitcer PL. The development of state and local injury surveillance systems. *J Safety Res.* 1987;18:191–198.
34. Thacker SB, Parrish RG, Trowbridge FL. A method to evaluate systems of epidemiologic surveillance. *World Health Stat Q.* 1988;41:11–18.
35. Sutter RW, Cochi SL, Brink EW, Sirotkin BI. Assessment of vital statistics and surveillance data for monitoring tetanus mortality, US 1979–1984. *Am J Epidemiol.* 1990;131:132–142.
36. Bureau of Labor Statistics. *Geographic Profile of Employment and Unemployment for 1985.* Washington, DC: Bureau of Labor Statistics; 1986.
37. Bureau of Labor Statistics. *Geographic Profile of Employment and Unemployment for 1986.* Washington, DC: Bureau of Labor Statistics; 1987.
38. US Department of Health and Human Services. *International Classification of Disease, 9th Revision, ICD-9.* Washington, DC: US Department of Health and Human Services; 1980. DHHS publication no. (PHS) 80-1260.
39. Bureau of Labor Statistics. *Recordkeeping Guidelines for Occupational Injuries and Illnesses.* Washington, DC: US Government Printing Office; 1986–491–543:54422.
40. Combs DL, Parrish RG, Ing R. Medical Examiner and Coroner Jurisdictions in the United States, U.S. Public Health Service, unpublished document, 1990.
41. Office of Management and Budget. *Standard Industrial Classification Manual, 1987.* Washington, DC.
42. Burnett CA, Crouse WE. Agreement of latest and longest occupation and industry as reported in the 1980 National Health Interview Survey. *J Occup Med.* 1989;31:339–346.
43. Conroy C, Russell JC. Medical examiner/coroner records: uses and limitations in occupational injury epidemiologic research. *J Forensic Sci.* 1990;35:932–937.

Occupational Health and Safety in Brazil

Howard Frumkin, MD, MPH, and Volney de M. Câmara, MD

ABSTRACT

Background. Brazil is the world's fifth largest and sixth most populous nation. Its economy is varied, with strong manufacturing, agriculture, mining, and service sectors. Therefore, a wide variety of workplace hazards confronts its work force. This paper describes Brazil's occupational safety and health regulatory scheme, workers' compensation system, plant-level practices, training, and data collection.

Methods. We reviewed and analyzed Brazilian regulatory legislation and government and non-governmental organization (NGO) activity in occupational safety and health, as well as the structure and function of the workers' compensation system. We also reviewed available data on injuries and diseases from major sources, including the now-defunct Instituto Nacional do Previdencia Social (INPS) and the workers' compensation scheme, Seguro de Acidente de Trabalho (SAT).

Results. The incidence of workplace injuries has decreased in recent years and is now reported to be about 5 per 100 workers per year. The case fatality rate has been constant at about 5 fatalities per 1000 injuries. Less than 6% of reported injuries are classified as "diseases." Brazil's rates are comparable to those of Mexico and Zimbabwe, and two to four times higher than in most industrialized countries.

Conclusions. Brazil has a high incidence of occupational injuries and diseases; these injuries and diseases are underreported; there is a large informal sector at special risk; and Brazil illustrates the disparity that exists in many countries between legislation on the books and legislation that is actually implemented. (*Am J Public Health.* 1991;81:1619–1624)

Background

History

Brazil is the world's fifth largest nation, with an area of 8.5 million km^2, and the sixth most populous, with approximately 150 million inhabitants.[1] It is a physically diverse nation, rich in natural resources. Until the 20th century, the economy depended principally on sugarcane, coffee, hardwood, and rubber production, and on gold and diamond mining. Because of fluctuating world markets for these products and dependence on a small number of export products at any time, Brazil's economy oscillated throughout the 19th century. During the 20th century, however, successive national governments have encouraged diversification. Industrialization has proceeded rapidly, and manufacturing now constitutes the largest portion of the gross domestic product. Tremendous economic growth has continued during most of the 20th century, especially during the 1960s and 1970s. However, serious economic and political problems persist, including marked income maldistribution, uncontrolled inflation, and one of the world's largest foreign debts.

Industry and Agriculture

The Brazilian economy includes well-developed mining, agriculture, manufacturing, and service sectors.

Rich mineral deposits support an extensive mining industry. There are major deposits of iron ore, bauxite, hematite, oil, and gas in many areas. Aluminum and manganese are mined in the Amazon region, chrome, magnesium, and quartz in Bahia, copper and lead in Bahia and Rio Grande do Sul; asbestos in Goias; and nickel in Goias and Minas Gerais. Brazil is self-sufficient in tin, zinc, and tungsten. Of particular interest is the gold found in the Amazonas region, which is mined by approximately 600 000 *garimpeiros* (gold-seekers) working outside the formal sector.[2] Mining employs about 4 million people nationally.[3]

Agricultural production has increased over recent decades as a result of increases in areas under cultivation, but the relative importance of agriculture in the nation's economy has declined with the growth of industry. Major crops include coffee (Brazil is the world's largest producer), soybeans, oranges, and sugarcane; the latter has increased in importance in recent years with use as a source of ethanol for fuel. Other crops include cassava, rice, cacao, and bananas. Brazil also has a sizable forestry industry, producing eucalyptus, pine, and a variety of hardwoods, and a significant fishery. Between 13 and 16 million people, slightly less than one third of the work force, work in the combined agriculture/forestry/fishing sector, representing a significant decline over recent decades.[2,4]

Brazil's manufacturing sector is a rapidly growing, vibrant part of the economy. Major products include steel, chemicals, petrochemicals, textiles, automobiles, pulp and paper, aircraft, electronics,

At the time of this study, Howard Frumkin was with the University of Pennsylvania School of Medicine. Volney de M. Câmara is with the Faculty of Medicine, Universidade Federal do Rio de Janeiro, Brazil. Both are affiliated with (INCLEN) the International Clinical Epidemiology Network.

Requests for reprints should be sent to Howard Frumkin, MD, MPH, Division of Environmental and Occupational Health, Emory University School of Public Health, 1599 Clifton Road, NE, Atlanta, GA 30329.

This paper was submitted to the journal August 10, 1990, and accepted with revisions May 6, 1991.

TABLE 1—Selected Substances Regulated Under the Consolidated Labor Law (ACGIH TLVs Provided for Comparison)

Substance	Tolerance Limit (ppm)	ACGIH TLV (ppm)	
Acetaldehyde	78	100	
Acrolein	...	0.1	
Ammonia	20	25	
Arsine	0.04	0.05	
Asbestos	...	0.2–2.0	C[a]
Benzene	8	10	C
1,3-Butadiene	780	10	
Cadmium	...	0.05 (mg/m³)	
Carbon monoxide	39	50	
Cellosolve	78	5	
Chromium compounds	...	0.005–0.5 (mg/m³)	
Ethylene oxide	39	1	C
Formaldehyde	1.6	1	C
Lead	0.1 (ceiling)	0.15	C
Methyl cellosolve	20	5	
Methylene chloride	156	50	C
Nickel carbonyl	0.04	0.05	C
Parathion	...	0.1 (mg/m³)	
Perchlorethylene	78	50	C
Vinyl chloride	156	5	C

Notes. C indicates a known or suspected carcinogen, for which the American Conference of Governmental Industrial Hygienists (ACGIH) recommends the lowest possible exposure, irrespective of threshold limit value (TLV). References: Norma Regulamentadoras Portaria 3214 de 8 de Junho de 1978; ACGIH, Threshold limit values and biological exposure indices for 1989–90, Cincinnati, 1989.
[a]Depends upon fiber type.

and footwear. The petrochemical industry is Latin America's largest. The manufacturing sector now employs about 25% of the work force, and continues to expand.[3,4] Civil construction employs large numbers of workers. The service sector is the fastest growing part of Brazil's economy, and now employs more than one third of the work force.

The greatest concentration of industry in Brazil is in the Southeast, centered around the cities of São Paulo, Rio de Janeiro, and Belo Horizonte. About one fifth of the nation's industry is in the South, and a very small portion is in the North.

The Brazilian Work Force

About 45% of Brazil's population, nearly 70 million people, are considered economically active, of whom about 60 million are employed. This includes about 20% of children aged 10 to 14, representing a significant pool of child labor.[5]

Most Brazilian workers are employed in small workplaces.[6,7] However, since no official data are gathered on employment by enterprise size, this trend is difficult to quantify.

Women constitute about 35% of Brazil's work force. Female workers tend to have completed more schooling than male workers. Overall, however, women are relatively more concentrated than men in the service sector and less concentrated in manufacturing, mining, agriculture, and construction, and they earn lower wages than men.[8]

Occupational Safety and Health Practice

History

Awareness of workplace accidents grew in the early years of the 20th century, along with larger concerns about the wages and living conditions of workers. Several state-level initiatives during the first two decades of the century addressed the problem of workplace accidents. A national law that regulated workplace safety conditions (Law No. 3724) was promulgated in 1919 and altered in 1934. During the decades of rapid industrialization that followed World War II, the number of reported accidents and their economic burden increased dramatically. The national government took action in 1968 by issuing Ordinance 32, which defined requirements for safety and health initiatives at the workplace level. Since then, further regulations have been promulgated on the federal level and in some states.

Regulation of Workplace Conditions

The legal framework for occupational health regulation is complex, ranging from the federal constitution (newly issued in 1988) to long-standing Ministry of Labor ordinances. Most occupational health functions have been housed in the Ministry of Labor, but with the reorganization of the health sector in 1988 into the Unified Health System (Sistema Unico de Saúde, or SUS), some responsibility was shifted from the Ministry of Labor to SUS. The new constitution empowers SUS to act in protecting worker health (Article 205, Sections II, VIII) but generally does not define how this is to be done. No specific piece of legislation defines a standard-setting process. The constitution does empower the Federal government to conduct workplace inspections (Article 21, Section XXIV). It remains to be seen how regulation and workplace inspection will be conducted by SUS under the new arrangement.

For the present, occupational health regulation continues to be an activity of the Ministry of Labor, pursuant to the Consolidated Labor Law of 1943 (Decreto-Lei 5452). In addition to federal activity, state and municipal governments issue some laws and regulations of their own. The Ministry of Labor has generally based its regulations on threshold limit values (TLVs) issued by the American Conference of Governmental Industrial Hygienists (ACGIH).[9,10] Of interest, the exposure limits are based on a 48-hour workweek, rather than the 40-hour workweek assumed in US limits. (The longer workweek remains common in Brazil, although the new constitution limited the workweek to 44 hours.) A list of regulated substances appears in Table 1. Several observations may be made. First, despite the equivalence of many standards (adjusted for differing workweeks), Brazil permits far higher exposure to certain substances, either through more permissive standards or through the complete absence of standards. Second, Brazilian standards based on older ACGIH TLVs have not been updated apace with ACGIH updates. For example, Brazil's methylene chloride standard is based on an older ACGIH TLV that, in the United States, was lowered in 1988. Third, there is far more attention to long-term outcomes, especially cancer, in ACGIH recommendations than in Brazilian regulations. Several carcinogens, such as chromium compounds, are not regulated in Brazil, and others are regulated less stringently.

Although some stringent regulations are in place, the Ministry of Labor has maintained virtually no workplace inspection capacity, and little enforcement has occurred. If a violation is discovered, a labor justice may impose a fine, but the maximum fine prescribed by law is only 10 times the minimum wage for a month (Chapter 5 of Law 5242 of 1943).

Occupational Safety and Health Service Delivery

Ordinance 32 of 1968, which was updated in 1983, requires companies to establish an Internal Committee for the Prevention of Accidents (CIPA). Specific requirements for this committee are based on the risk level and on the number of employees.[11] The main objectives are the prevention of accidents and the monitoring of safety and health conditions in the plant. Although the committees must contain an equal number of labor and management representatives, the chairman is a management representative. Moreover, although Ordinance 32 protects CIPA labor representatives from discharge because of their committee activity, anecdotal evidence suggests that this protection is sometimes violated. These factors compromise the independence and effectiveness of the workplace committees.

In addition to workplace safety and health administration, professional staff requirements are defined by law. A 1972 ordinance (No. 3237) mandated specialized hygiene and occupational medicine services for enterprises according to the number of employees and risk level. For example, a high-risk firm with 1500 employees is required to maintain one occupational physician, one aide, one safety inspector, and one occupational engineer. The ordinance defines the occupational medical role broadly as including preemployment and periodic medical examinations, worker education, surveillance activity, etiologic study, and committee work at the plant level.

Workers' Compensation

The workers' compensation function was fully integrated into the nation's larger social welfare system in 1967 by Law 5316, which required all employers to insure their employees through what was then the National Institute for Social Welfare (Instituto Nacional do Previdencia Social, or INPS). Under the restructuring introduced by President Collor de Mello in 1990, the INPS was dissolved. All its social welfare functions, including workers' compensation, were transferred to the Ministry of Labor in the newly formed Instituto Nacional de Seguridade Social (INSS). Although some changes may occur, it is likely that the INSS compensation scheme will preserve many of the features it developed under INPS between 1967 and 1990.[5]

The workers' compensation scheme is called Seguro de Acidentes de Trabalho (SAT). In case of an occupational injury or disease, this scheme offers substantial benefits above ordinary medical coverage. These include medical costs, full wage replacement, compensatory payments for loss of function, rehabilitation costs, and death benefits.

The INSS is funded by equal contributions from employees and employers, each composing 10% of the payroll. Because INSS also acts as Brazil's workers' compensation insurance carrier, it collects extra funds to support this function. Payments are collected from employers according to each industry's risk profiie. Premiums range from 2.5% of the employees' salaries in high-risk enterprises to 0.4% in low-risk enterprises. Risk in this context is evaluated on the basis of lost time accidents (of more than 15 days), a point to which we return below.

When an occupational injury or illness occurs, the employer is required to report it to the INSS within 24 hours. The report automatically triggers SAT coverage, which pays for medical care. The employer continues to pay full wages for the first 15 days of lost time, at which point the SAT takes over with continued full wage replacement. If a company fails to report an injury or illness, claiming that it is not occupational, then a worker may take the case to labor court.

There are many failures of this system.[12] First, the SAT scheme includes only about 25% of the working population, since only about half of workers are covered by the INSS, and only half of those work for companies required to pay into SAT. Civil servants, physicians, and self-employed people, for example, would typically be covered by INSS but not by SAT. Rural workers, who constitute close to 30% of Brazil's work force, are also not covered by SAT. Moreover, companies have an incentive not to report injuries and illnesses, mostly a matter of public image rather than finances. This is especially true of injuries that result in less than 15 lost workdays, since the SAT offers wage replacement only after that interval. In addition, physicians often fail to diagnose occupational illnesses correctly, and workers often are unaware of procedures for submitting injury reports and gaining compensation. For all these reasons, the SAT fails to provide compensation to many workers who become injured or ill at work. The extent of this problem is unknown.

Of additional importance, SAT reports are the sole source of occupational injury and illness data in most of Brazil. Some states, such as Rio de Janeiro, require reporting of certain occupational conditions, but this is a variable and incomplete resource. Therefore, the failures of the workers' compensation system translate into failures of occupational health and safety data as well.

Labor Union Activity

Approximately 22% of Brazilian workers belong to labor unions.[13] The unions have been relatively inactive in occupational health and safety because of restrictions imposed by the military government until 1985 and because of more pressing economic concerns. However, some labor unions have recently increased their activity in the field, perhaps supported by growing awareness of environmental devastation in some parts of the country with consequent health costs. In 1979 the two labor union confederations, Central Unica de Trabalhadores (CUT) and Confederação Geral de Trabalhadores (CGT), formed the Interunion Department for the Study of Health and Work Environment (Departamento Intersindical de Estudos de Saude e dos Ambientes do Trabalho, or DIESAT). DIESAT advises local unions on safety and health problems, publishes a monthly newsletter, offers seminars, and participates in investigations. The increasing union activity has been evident at the state and national level. In 1980, for example, the "Sindicatos dos Trabalhadores Rurais de Campos-Rio de Janeiro" won a national prohibition of the use of organomercurial fungicides. In 1988, pressure by the keypunchers union, Associação dos Profissionais de Processamento de Dados (APPD), resulted in recognition by the state government of Rio de Janeiro of tenosynovitis of the fingers as a reportable occupational disease among keypunch operators.

Injury and Disease Reporting

In Brazil, workplace injuries are defined as "all events that happen during the practice of work directly or indirectly causing physical lesion, functional disturbance or illness, and resulting in death or

TABLE 2—Designated Causes of Occupational Diseases According to Decreto 79.037, December 1976

Arsenic
Asbestos
Benzene, its homologues, and nitro and amino derivatives
Beryllium
Cadmium
Carbon disulfide
Carbon monoxide
Chromium and its compounds
Fluorine, chlorine, iodine, and bromine
Halogenated hydrocarbons
Iodine
Ionizing radiation
Lead
Manganese
Mercury
Phosphorus
Pressure, atmospheric
Silica dust
Tar, asphalt, paraffin, and related materials

loss of the capacity to work, whether total or partial, temporary or permanent."[14] These include injuries at work or while commuting, officially designated occupational diseases, and other diseases found to be work-related. The list of designated causes of occupational diseases (see Table 2) was set forth in December 1976 by Decreto 79.037. For some diseases, Decreto 79.037 specifies particular operations and techniques that are considered pathogenic. This may result in underrecognition of occupational disease; if an operation that causes disease was omitted by the ordinance or was introduced after the ordinance was written, it is not acknowledged as a cause of occupational disease.

The report of an injury is the responsibility of the employer. After the employer has been informed of the injury, the Ministry of Labor must be notified within 24 hours and given a description of the injury.[11] These data are tabulated, but very little further analysis is performed. There is no systematic effort to identify hazardous enterprises or to apply these data to enforcement or policy decisions.

One source of error in injury statistics derives from the fact that commuting injuries are notifiable. When a worker is injured over the weekend, he or she might report that injury as a commuting injury upon arriving at work Monday morning. This phenomenon is apparently common in some companies, although difficult to quantify. It may partially explain the peak in injuries seen on Mondays (data not shown).

Training

Medical and nursing curricula in Brazil vary in their occupational safety and health content. In major universities in Rio de Janeiro, São Paulo, and Minas Gerais, considerable time is devoted to occupational safety and health. For example, the medical curriculum at the Federal University of Rio de Janeiro includes 126 contact hours in occupational medicine during the third year. Other universities have considerably less instruction in occupational safety and health.

Specialty training in occupational safety and health was greatly promoted by the 1972 ordinance that requires occupational physicians, industrial hygienists, and occupational health nurses at plants. These functions were initially well salaried in a nation that suffers from professional unemployment, so large numbers of physicians and nurses sought certification. Within several years, universities responded to the increased demand by offering large numbers of training programs. The result was a large complement of occupational safety and health professionals by the late 1970s; over 20 000 occupational physicians were trained during that decade.[15] Presently, about 10 universities throughout the nation offer specialty training in occupational medicine, occupational nursing, and industrial hygiene. Certification in occupational medicine or nursing requires 360 hours of instruction over 1 year, usually taken as an evening course. Upon completion of the course, the Ministry of Labor certifies graduates, which qualifies them to fill the role defined by law. This function may be taken over by the Ministry of Education in coming years.

Available Data on Occupational Injuries and Diseases

The data in this report are derived from the population covered by the SAT scheme. Although the Ministry of Labor must be notified of all workplace injuries, only those injuries resulting in at least one lost workday are reported. In contrast, most other countries' data are based on 3-day lost work time injuries. The numerator data reported here were obtained from the INPS, and the denominator data were obtained from the Brazilian Institute for Geography and Statistics, or IBGE.[8] A foundation based in the Ministry of Labor, FUNDACENTRO, also publishes workplace injury reports in its bulletins, again based on INPS data.

Figure 1 shows the occupational injury incidence rate (including both occupational diseases and transit injuries) per 100 employees. Despite the deficiencies in reporting, the data indicate that Brazil has a high incidence of occupational injuries. This rate increased until 1972 and decreased thereafter. At least two explanations may be offered for the change: improving working conditions or less complete reporting of injuries.

In support of the first explanation, at least two developments may be cited. One is the legislation that mandated plant-level occupational health services. Second, increased training in occupational safety and health at the university level might have contributed to a decrease in workplace injuries.

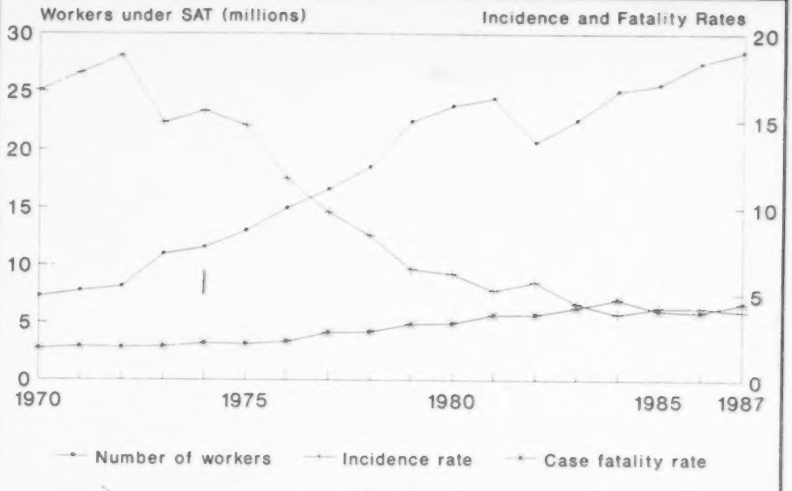
In support of the second explanation, several observations are relevant. First, as noted above, most of the Brazilian working population is employed in small enterprises that were not required to establish Internal Committees and Occupational Health Services. Second, a major change in Brazil's insurance legislation occurred in 1975. Until then, the employer was required to pay a worker's salary on the day of a workplace injury, while the INPS assumed this role on the second day. After 1975, employers were required to pay the injured worker's salary until the 15th day after the injury. This change was a powerful incentive for companies not to report injuries. Unfortunately, the Ministry of Labor has not monitored companies' compliance in recording and notifying injuries, so the extent of underreporting is unknown.

Several other observations suggest that the reported data are underestimates. First, the SAT data cover only about 25% of the work force, excluding especially hazardous sectors such as agriculture. Second, the number of reported occupational diseases (which are included in injury data) is very low. Physicians in Brazil are still not familiar with the relationship between work and health,[12] and specialized diagnostic services such as toxicological laboratories are rare. Finally, the apparent decrease in occupational injury incidence did not correspond to a similar decrease in the occupational fatality rate. In Brazil, as in other developing countries, mortality records are usually more reliable than morbidity and other data. Figure 1 shows that the case fatality rate increased from 1970 through 1987. The likely explanation is that fatal injuries continued to be notified, while nonfatal injuries did not. Although international comparisons are

difficult, it appears from available data that Brazil's workplace fatality rate is significantly above that of other countries, as shown in Table 3.

Figure 2 shows the proportion of occupational diseases and transit injuries included in the total number of injuries. Reported cases of occupational disease have never exceeded 1% of the total, probably reflecting significant underreporting of disease. On the other hand, there has been an increasing trend in transit injuries, reflecting better reporting and/or increasing risk in a nation with increasingly congested and dangerous roads.

The rate of occupational disease notification in Brazil was estimated during a study of organomercury fungicide use by sugarcane workers in the district of Campos, State of Rio de Janeiro.[16] For the year 1980, 1456 occupational injuries were registered, including one occupational disease. In this same year, 30 sugarcane workers were submitted to a clinical and laboratory examination, which revealed 21 cases of mercury poisoning with blood levels of mercury as high as 0.25 ppm. None of the 21 affected workers was among the 1456 reported cases, suggesting widespread underreporting.



Sources: INPS, IBGE

FIGURE 1—Occupational injury incidence (per 100 workers) and case fatality (per 1000 injuries) rates, Brazil, 1970–1987.

Conclusions

Several conclusions emerge from this review of occupational health and safety in Brazil. First, there is a high incidence of occupational injuries and diseases. This is common among developing countries, often to a worse degree than in developed nations. Brazil illustrates the tension that may exist between economic development and attention to health concerns. In nations with crushing foreign debts, like Brazil, the economic crisis may eclipse less immediate and pressing concerns such as occupational safety and health. Constituencies that might promote safety and health measures, such as labor unions and the health sector, are forced in this situation to direct their efforts to other issues. This tension exists in all developing nations, and may be expected to continue in the foreseeable future. It highlights the importance of greater attention to occupational health in Brazil, and of vigorous attempts to prevent injuries and illnesses.

Second, there is evidence of serious underreporting of occupational injuries and illnesses, worse with illnesses than with injuries. Among the most important reasons are the exclusion of a majority of the work force from the only systematic data collection system, reliance on employers for notification despite incentives against notifying, and underdiagnosis of occupational diseases by physicians. On the other hand, two factors may increase reporting relative to other nations: the inclusion of injuries that occur while commuting and the reporting of injuries that result in as little as one lost workday. In any case, more consistent, standardized injury and disease surveillance is clearly required.

A third conclusion relates to the informal sector of Brazil's economy. As is the case in many developing nations, a large portion of Brazil's work force is marginalized and excluded from services. Examples include the street vendors who are ubiquitous on the streets of major cities, agricultural workers, and the gold-seekers in the Amazon basin. These groups may be at special risk of occupational injury and disease, and special efforts must be made to extend surveillance and prevention programs to them.

Brazil also illustrates the disparity that may exist between legislation "on the books" and reality. It has a relatively complete set of laws that relate to occupational safety and health, but enforcement is erratic and generally absent. As a result, hazardous conditions persist in many workplaces. Similarly, although national legislation triggered training of a large number of occupational health professionals, especially physicians, this group has not had the impact on workplace conditions that might have been hoped. These observations suggest that vigorous action by regulators, labor unions, and professionals may be necessary to achieve real change.

TABLE 3—Occupational Fatality Rates in Several Countries

Country	Fatality Rate Per 1000 Employees
Zimbabwe	.217
Brazil	.220
Cuba	.102
Mexico	.240
United States	.068[a]
Japan	.025[a]
Spain	.127
Holland	.016
Poland	.109

Note. Mexico data are for 1986, Brazil and US data are for 1987, Zimbabwe data are for 1989, and all other countries' data are for 1988. *Source:* International Labour Office. *Year Book of Labour Statistics, 1989–90.* Geneva, Switzerland: ILO; 1990: Table 30.
[a]Rates for Japan and the United States are given as fatalities per 1 000 000 person hours worked. These rates were multiplied by 2.5 to convert to fatalities per 1000 persons.

An interesting and unusual feature of Brazil's occupational health landscape is the integration of almost all occupational health services into the health sector. Historically, the provision of social services has been tied to work. Almost 25 years ago, the workers' compensation function was integrated into the INPS, although

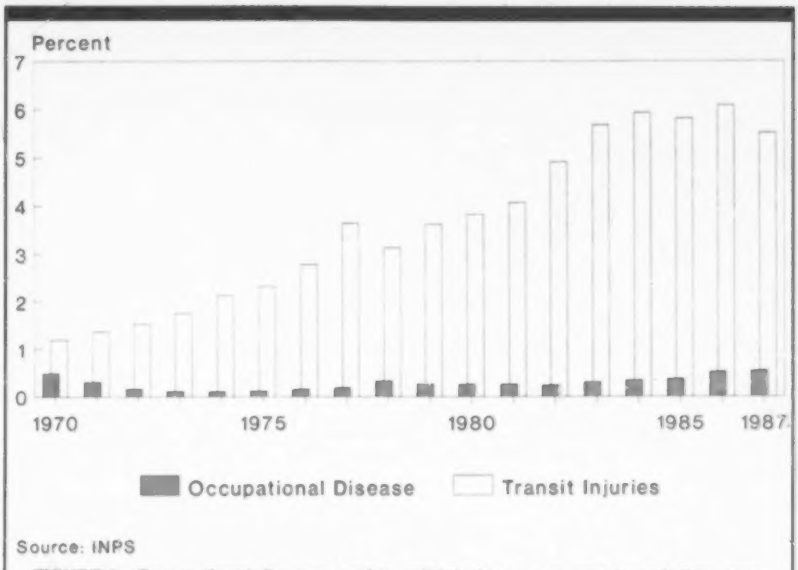


FIGURE 2—Occupational diseases and transit injuries as a proportion of all occupational injuries, Brazil, 1970–1987.

separate sites for delivering occupational medical services were maintained. In recent years, especially with the new constitution, other functions such as workplace inspection seem to be moving from the Ministry of Labor to the Ministry of Health, although some reversal of this trend occurred in 1990 under the new government. The effects of this change remain to be seen, but if successful, it offers some promise of decreased intersectoral disputes and a more efficient government role in occupational safety and health. □

References

1. US Department of State. *Brazil: Background Notes.* Washington DC: Department of State; 1990.
2. Couto RC, Câmara VM, Sabrosa P. Intoxicação mercurial: resultados preliminares em duas áreas garimpeiras. *Cadernos de Saúde Coletiva.* 1988;4:301–315.
3. Brazil. In: *Encyclopaedia Britannica.* 15th ed. Chicago, Ill: Encyclopaedia Britannica Inc; 1987:15:188–212.
4. International Labour Office. *Year Book of Labour Statistics, 1989–90.* Geneva: ILO; 1990.
5. Tambellini AT. *Política nacional de saude do trabalhador: alánises e perspectivas.* Rio de Janeiro: FIOCRUZ; 1986. Mimeo.
6. Possas C. *Saúde e Trabalho: A Crise da Previdência Social.* 2nd ed. São Paulo, Brazil: Ed. Hucitec; 1983.
7. Mendes R. *Importancia das Pequenas Empresas Industriais no Problema Dos Acidentes de Trabalho.* São Paulo, Brazil: Tese de Mestrado, Universidade do São Paulo; 1975.
8. Instituto Brasileiro de Geografia e Estatística. *Pesquisa Nacional por Amostra de Domicilios*—1987. Rio de Janeiro, Brazil: Instituto Brasileiro de Geografia e Estatística; 1988.
9. Ministério do Trabalho. *Curso de Medicina do Trabalho.* Sao Paulo: Fundacentro; 1979.
10. Ribeiro HP, Lacaz FA de C. *De Que Adoecem e Morrem os Trabalhadores.* São Paulo: Ed. Diesat; 1984.
11. Manual de Legislação Atlas. *Segurança e medicina do Trabalho, 9ª Ediçao.* Rio de Janeiro, Brazil: Ed. Atlas; 1984.
12. Mendes R. Perfil de morbi-mortalidade dos trabalhadores no Brasil. Presented to SBPC, Porto Alegre; 1990. Mimeo.
13. US Department of Labor, Bureau of International Labor Affairs. *Foreign Labor Trends: Brazil.* Washington, DC: US Department of Labor. FLT 90-31.
14. Diario Oficial do Brasil, May 3 1944.
15. Relatório da I Reunião Nacional sobre Ensino e Pesquisa em Saúde Pública, Relatório Final. In: *Ensino da Saúde Pública, Medicina Preventiva e Social do Brasil.* Rio de Janeiro, Brazil: ABRASCO; 1984.
16. Silva MRC, Câmara VM, Senone M. *Otimização ergonômica dos tratos culturais da lavoura de cana-de-açucar.* 4° Relatório Técnico, Fundação Getulio Vargas, UFRJ, Rio de Janeiro; 1980. Mimeo.
17. Norma Regulamentadoras Portaria 3214 de 8 de Junho de 1978.
18. American Conference of Governmental Industrial Hygienists. Threshold limit values and biological exposure indices for 1989–90, Cincinnati, 1989.

An Evaluation of the Impact of Maternity Care Coordination on Medicaid Birth Outcomes in North Carolina

Paul A. Buescher, PhD, Marcia S. Roth, MPH, Dennis Williams, BA, and Carolyn M. Goforth, BSN

ABSTRACT

Background. Care coordination is an important component of the enhanced prenatal care services provided under the recent expansions of the Medicaid program. The effect of maternity care coordination services on birth outcomes in North Carolina was assessed by comparing women on Medicaid who did and did not receive these services.

Methods. Health program data files, including Medicaid claims paid for maternity care coordination, were linked to 1988 and 1989 live birth certificates. Simple comparisons of percentages and rates were supplemented by a logistic regression analysis.

Results. Among women on Medicaid who did not receive maternity care coordination services, the low birth weight rate was 21% higher, the very low birth weight rate was 62% higher, and the infant mortality rate was 23% higher than among women on Medicaid who did receive such services. It was estimated that, for each $1.00 spent on maternity care coordination, Medicaid saved $2.02 in medical costs for newborns up to 60 days of age. Among the women who did receive maternity care coordination, those receiving it for 3 or more months had better outcomes than those receiving it for less than 3 months.

Conclusions. These results suggest that maternity care coordination can be effective in reducing low birth weight, infant mortality, and newborn medical care costs among babies born to women in poverty. (*Am J Public Health.* 1991;81:1625–1629)

Introduction

Factors associated with infant mortality and morbidity and with handicapping conditions in infants and young children are numerous and complex. It has been documented repeatedly, however, that the single most important factor contributing to infant death and disability is low birth weight.[1,2] Avoiding unintended and high-risk pregnancies is an important first line of prevention.[3] Experience proves that the incidence of low birth weight can also be reduced through the use of early, continuous, and comprehensive prenatal care.[4]

In October 1987, in response to concerns over North Carolina's high infant mortality rate, the state Medicaid program was expanded to improve access to health care and support services for low-income pregnant women and young children. This expansion, referred to in North Carolina as the Baby Love Program, was developed and implemented by the Division of Maternal and Child Health and the Division of Medical Assistance, in cooperation with the Office of Rural Health and Resource Development. Two key elements of the Baby Love Program were expansion of Medicaid eligibility to 100% of the federal poverty level and reimbursement for maternity care coordination (case management) services for Medicaid-eligible pregnant women.

Maternity care coordination is the cornerstone of the Baby Love Program in that it is aimed directly at eliminating barriers to client use of services. Maternity care coordinators help eligible women receive services that address not only medical but also nutritional, psychosocial, and resource needs. Coordinators also provide pregnant women with social and emotional support, which may lead to stress reduction and the adoption of healthful behaviors during pregnancy. In North Carolina, all pregnant women certified for Medicaid are eligible for maternity care coordination services. This entire population is considered to be at high risk for poor pregnancy outcomes, owing to either medical or psychosocial risk factors, or a combination of the two.

Previous studies have shown that comprehensive prenatal care, which includes various nonmedical support services, improves birth outcomes among women in poverty.[4,5] The purpose of this evaluation is to determine if the receipt of maternity care coordination services by Medicaid-eligible clients improves birth outcomes. Maternity care coordination, which began in late 1987, was introduced as a statewide program. Because the program was not piloted in a limited geographic area of the state, it was important to conduct the evaluation before the statewide system was fully developed to ensure that a comparison group of nonrecipients could be obtained.

The period of evaluation was 1988 and 1989, the first 2 years of program im-

Paul A. Buescher is with the State Center for Health and Environmental Statistics, and Marcia S. Roth is with the Division of Maternal and Child Health both in the Department of Environment, Health, and Natural Resources, Raleigh, NC. Dennis Williams and Carolyn Goforth are with the Division of Medical Assistance, Department of Human Resources in Raleigh, NC.

Requests for reprints should be sent to Paul A. Buescher, PhD, State Center for Health and Environmental Statistics, PO Box 27687, Raleigh, NC 27611-7687.

This paper was submitted to the journal September 4, 1990, and accepted with revisions May 28, 1991.

plementation. During this time, more than 15 000 women received maternity care coordination, two thirds of whom were served in 1989. For the 2-year period, the mean month of pregnancy in which maternity care coordination began was 5.6. In 1988, the start-up year, the mean month of service initiation was 6.0; in 1989, this had improved to 5.4.

Methods

North Carolina resident live births to women on Medicaid during calendar years 1988 and 1989 were examined to compare birth outcomes of those women who received maternity care coordination services with the outcomes of those who did not. The birth outcomes compared were low birth weight (less than 2500 g), very low birth weight (less than 1500 g), infant mortality, and newborn medical care costs (which are, to some degree, a surrogate for infant morbidity). The effect of maternity care coordination on prenatal participation in the Special Supplemental Food Program for Women, Infants and Children (WIC) was also examined.

To carry out this evaluation, we linked a number of data files to 1988 and 1989 birth certificate records. Using the baby's name, date of birth, hospital of birth, and other information, we matched birth records to newborn hospital claims paid by Medicaid to identify "Medicaid" births and achieved a matching rate of over 95%. Then, using the baby's Medicaid ID number retained from this step, we extracted all claims paid for any service (inpatient, outpatient, medications, etc.) *beginning* within 60 days of birth and we summarized the costs. If a hospitalization began within 60 days of age but extended beyond that time all costs for the hospitalization were counted. These costs were examined only for services beginning within 60 days of birth because maternity care coordination would be expected to have its major impact on early infant morbidity. Also, the babies of women receiving maternity care coordination are automatically eligible to remain on Medicaid for 60 days after birth; thus, there is no problem of attrition within this group due to discontinued eligibility.

Using the mother's name and date of birth, we matched birth records to claims paid for maternity care coordination to identify births for which the service was received. Total amount paid for care coordination was retained from this step. Then we summarized prenatal visit records from the public health department client information system in North Carolina and, again using mother's name and date of birth, matched them to the birth records to identify those births that received prenatal clinic care in health departments. And we used mother's name and date of birth to match births with WIC records to identify women receiving this nutritional service during the prenatal period. Finally, infant death records were linked to the 1988 and 1989 birth cohorts as part of routine vital records processing, and this allowed us to assess infant mortality rates.

In matching the maternity care coordination, health department, and WIC records to the 1988 and 1989 births, we achieved a matching rate of about 90%. Incomplete matching of the maternity care coordination records has resulted in some maternity care coordination recipients being counted as nonrecipients. To the extent that maternity care coordination does improve birth outcomes, this incomplete matching has caused the findings shown here to understate the true difference between the two groups.

Within the group of Medicaid births identified by a matching Medicaid newborn hospital claim, we compared birth outcomes between women with and without maternity care coordination. Multiple births (twins etc.) and births with no prenatal care, the latter of which would all fall into the not-care-coordinated group and thus bias the results, were excluded from the analysis to increase the validity of the comparisons. Simple comparisons of percentages and rates were supplemented by a logistic regression analysis. This analysis assessed the effect of maternity care coordination on birth outcomes while statistically controlling for differences between the two groups on other measurable risk factors.

We also compared outcomes for women on Medicaid with different lengths of maternity care coordination, as defined by months of participation (less than 3 months vs 3 months or more). Because there could be some bias in the results by length of participation, given that a woman who delivers prematurely will thereby have a shorter period of care coordination participation, we compared low birth weight rates between these groups for full-term births (at least 37 weeks gestation). As another way to control for this gestational age bias, we examined birth outcomes by length of maternity care coordination participation in weeks as a *percent* of length of gestation in weeks (less than 25%, 26 to 50%, and more than 50%).

Previous studies have shown that prenatal care provided in public health clinics is very effective in reducing adverse birth outcomes among women in poverty.[4,5] Therefore, it might be argued that women receiving maternity care coordination would have had better birth outcomes simply because most of them received prenatal care in public health departments, particularly in the first years of the program; whereas many of the women not receiving maternity care coordination received their prenatal care from another provider. To control for this potential source-of-prenatal-care bias, we broke out the basic results of this study separately for public health department patients and for patients of other prenatal care providers. We also used source of prenatal care as a control variable in the logistic regression analysis.

Results

Measures in the top part of Table 1 reveal that, among women on Medicaid not receiving maternity care coordination services, the low birth weight rate was 21% higher, the very low birth weight rate was 62% higher, and the infant mortality rate was 23% higher than among women on Medicaid who did receive such services. Further, the average cost of medical care for the baby in the first 60 days of life was $277 higher for infants of women not receiving maternity care coordination services, while prenatal WIC participation was substantially lower for this group.

The results in Table 1 are also shown separately for women receiving prenatal clinic care in public health departments and for those receiving prenatal care from other providers. These results are similar to those for the total: the women receiving maternity care coordination had consistently better outcomes in both prenatal care groups. Therefore, the better birth outcomes for women receiving maternity care coordination are not just a result of most of them receiving prenatal care in the public health setting.

For the total group of 15 526 women receiving maternity care coordination, Medicaid paid an average of $137 for the care coordination services. Dividing the difference in average newborn costs of $277 by this $137 results in an estimated $2.02 saved by Medicaid in all medical costs for services beginning within 60 days of age for each $1.00 spent on maternity care coordination. An average net savings of $140 ($277−$137) for each of the 15 526 women receiving care coordination trans-

lates into a total estimated savings of $2,174,000 for 1988 to 1989.

Table 2 shows that the women on Medicaid who received maternity care coordination services were at somewhat *higher* risk for several of the maternal characteristics often associated with low birth weight and infant death (characteristics derived from the birth certificates). Therefore, their better birth outcomes shown in Table 1 do not appear to be owing to favorable status on these measured risk factors. This slightly higher demographic risk may be related to the fact that some health departments initially targeted maternity care coordination to teens and other high-risk groups because insufficient resources were available to provide this service to all Medicaid-eligible women. Yet the pregnant women who received maternity care coordination had better birth outcomes despite higher levels of these other risk factors.

Table 3 presents the results of a logistic regression analysis that compared the independent effects of various risk factors on low birth weight and infant mortality. Race, age of mother, and medical risk factors during the pregnancy were the strongest predictors of adverse outcomes. After controlling for all the other risk factors shown in Table 3, including source of prenatal care, women not receiving maternity care coordination services were found to be 1.22 times as likely to have had a birth under 2500 g as women receiving the service, 1.48 times as likely to have had a birth under 1500 g, and 1.20 times as likely to have had an infant death. Both low birth weight odds ratios have 95% confidence intervals with the lower limit above 1.00.

So far in this analysis, maternity care coordination has been treated as present or absent, regardless of the length of care coordination for those receiving the service. In Table 4, the women receiving care coordination are divided into two groups: those beginning care coordination 3 or more months before giving birth and those receiving care coordination for a shorter period of time. For each measure in Table 4, the women who received care coordination longer had better outcomes.

The potential problem of bias due to preterm delivery shortening the length of service participation was addressed in two ways. First, we compared only full-term births. Among full-term births, women with care coordination for 3 or more months had a low birth weight rate of 3.81% compared with a rate of 5.07% for those receiving care coordination for less than 3 months ($P < .001$). Second, we compared birth outcomes by percent of the pregnancy for which maternity care coordination was received. Table 5 shows these results. Women who received maternity care coordination for more than 50% of their pregnancy had substantially lower rates of low birth weight, very low birth weight, and infant mortality. These results are further evidence that maternity care coordination services for pregnant women on Medi-

TABLE 1—Outcome Measures for Singleton Medicaid Live Births by Presence of Maternity Care Coordination: North Carolina, 1988–1989 (Births with No Prenatal Care Are Excluded)

Outcome Measures	Received Care Coordination	Did Not Receive Care Coordination	*P* Value for Difference
Total			
Percent under 2500 g	8.67	10.50	<.0001
Percent under 1500 g	1.23	1.99	<.0001
Infant deaths per 1000 live births	9.9	12.2	.02
Percent on prenatal WIC	90.9	66.8	<.0001
Average amount paid by Medicaid for newborn care beginning within 60 days of age	$1694	$1971	<.0001
Number of births	15 528	34 463	
Prenatal Care at Public Health Department			
Percent under 2500 g	8.61	9.95	.0001
Percent under 1500 g	1.16	1.71	.0001
Infant deaths per 1000 live births	9.7	12.5	.03
Percent on prenatal WIC	92.8	84.6	<.0001
Average amount paid by Medicaid for newborn care beginning within 60 days of age	$1674	$1872	<.0001
Number of births	12 842	14 812	
Prenatal Care Not at Public Health Department			
Percent under 2500 g	8.95	10.91	.001
Percent under 1500 g	1.57	2.21	.02
Infant deaths per 1000 live births	10.8	12.0	.58
Percent on prenatal WIC	81.9	53.4	<.0001
Average amount paid by Medicaid for newborn care beginning within 60 days of age	$1790	$2046	.001
Number of births	2684	19 651	

Note. WIC = (Special Supplemental Food Program for) Women, Infants and Children.

TABLE 2—Percentages of Singleton Medicaid Live Births with Selected Maternal Risk Factors by Presence of Maternity Care Coordination: North Carolina, 1988–1989 (Births with No Prenatal Care Are Excluded)

Risk Factors	Received Care Coordination	Did Not Receive Care Coordination	*P* Value for Difference
Black[a]	60.8	58.1	<.0001
Unmarried	66.4	61.2	<.0001
Under age 18	15.6	13.2	<.0001
Education less than 12 years	48.5	44.2	<.0001
Previous fetal or live-born death	24.1	25.0	.06
Prenatal care less than adequate in quantity[b]	49.8	49.2	.19
Mother smoked	30.4	31.2	.07
Medical risk factors this pregnancy	24.4	24.2	.58

[a]Data shown are actually for births for all races other than White, about 84% of which are Black in North Carolina.
[b]Based on intermediate plus inadequate categories of the Kessner Index of the quantitative adequacy of prenatal care use.[9] Missing values for this and the other variables in Table 2 are less than 1% of the total and are excluded from all calculations.

TABLE 3—Estimated Relative Odds of Low Birth Weight and Infant Death with Presence of Selected Risk Factors (Logistic Regression): 1988–1989 Singleton Medicaid Live Births in North Carolina (Births with No Prenatal Care Are Excluded)

Risk Factors	<2500 g	<1500 g	Infant Death
Black	1.66 (1.78, 1.54)	1.69 (1.99, 1.43)	1.14 (1.39, 0.94)
Unmarried	1.15 (1.23, 1.07)	1.10 (1.30, 0.94)	0.99 (1.20, 0.81)
Under age 18	1.29 (1.42, 1.18)	1.31 (1.63, 1.05)	1.23 (1.59, 0.95)
Education less than 12 years	1.12 (1.20, 1.05)	0.98 (1.14, 0.84)	1.09 (1.31, 0.90)
Previous fetal or live-born death	1.26 (1.35, 1.18)	1.74 (2.01, 1.50)	1.24 (1.50, 1.03)
Prenatal care less than adequate in quantity	1.13 (1.20, 1.06)	0.94 (1.07, 0.81)	1.05 (1.24, 0.88)
Mother smoked	1.73 (1.84, 1.62)	1.17 (1.37, 1.01)	1.22 (1.47, 1.02)
Medical risk factors this pregnancy	2.14 (2.27, 2.01)	2.94 (3.37, 2.56)	1.79 (2.13, 1.50)
Received prenatal care outside public health department	1.12 (1.19, 1.05)	1.33 (1.54, 1.15)	1.02 (1.23, 0.85)
Did not receive care coordination services	1.22 (1.32, 1.14)	1.48 (1.76, 1.24)	1.20 (1.47, 0.98)
Number of observations	49 260	49 260	49 288

Note. The 95% confidence interval limits for the odds ratios are shown in parentheses.

TABLE 4—Outcome Measures for Singleton Medicaid Live Births by Length of Maternity Care Coordination: North Carolina, 1988–1989 (Births with No Prenatal Care Are Excluded)

Outcome Measures	Care Coordination 3 or More Months	Care Coordination Less Than 3 Months	*P* Value for Difference
Percent under 2500 g	6.47	10.53	<.0001
Percent under 1500 g	0.70	1.71	<.0001
Infant deaths per 1000 live births	8.7	11.6	.08
Percent on prenatal WIC	92.3	80.0	<.0001
Average amount paid by Medicaid for newborn care beginning within 60 days of age	$1483	$1879	<.0001
Number of births	7680	7846	

TABLE 5—Outcome Measures for Singleton Medicaid Live Births by Percentage of Pregnancy for Which Maternity Care Coordination Was Received: North Carolina, 1988–1989 (Births with No Prenatal Care Are Excluded)

Outcome Measures	Less Than 26 Pct.	26–50 Pct.	More Than 50 Pct.	*P* Value for Difference (Col. 1 vs Col. 3)
Percent under 2500 g	9.88	6.25	6.44	<.0001
Percent under 1500 g	1.44	1.20	0.72	.001
Infant deaths per 1000 live births	10.1	10.6	8.0	.29
Number of births	6587	4908	4031	

caid in North Carolina are effective in reducing adverse birth outcomes.

Discussion

These results strongly suggest that maternity care coordination did improve the outcomes of live births to women on Medicaid in 1988 and 1989, the first 2 years of the Baby Love Program in North Carolina. The apparently large impact of maternity care coordination on very low birth weight is especially important. Even modest reductions in this relatively rare event can make prevention programs cost-effective, given the very high rates of morbidity and the use of medical care among these infants.

The results by degree of maternity care coordination (Tables 4 and 5) provide additional evidence that care coordination improved birth outcomes among these women in 1988 and 1989. Longer participation was associated with better birth outcomes, even after controlling for potential gestational age bias.

Women who receive maternity care coordination are encouraged to obtain a variety of services for which they may be eligible, such as Aid to Families with Dependent Children, job training, social work, transportation, food stamps, and housing assistance. Many of these services are not directly reimbursable by Medicaid. The maternity care coordinator acts as an advocate, assisting the pregnant woman in navigating a complex service system with often confusing application procedures so that such services may be obtained. The visit schedule is determined by the client's need and may be much more frequent than the minimum required level of one encounter per month.

One example of the impact of this assistance was prenatal WIC participation. We found that those women on Medicaid with maternity care coordination were significantly more likely to receive this important nutritional service than those without it. In addition, data for 1988 and 1989 from the North Carolina Health Services Information System show that, among public health department patients, women receiving maternity care coordination were more likely to have had a postpartum family planning exam (68% vs 39%) and to have had their newborns receiving well baby care (65% vs 25%) and WIC (77% vs 36%) than women not receiving care coordination.

In a study such as this, in which it was not possible to randomize patients into treatment and control groups, there is always the possibility of selection bias. Women who would have had better birth outcomes anyway may have been drawn for various reasons into the maternity care coordination group. Conversely, women who have experienced previous adverse pregnancy outcomes may have been channeled into care coordination because of their high-risk status. We have done our best to statistically control for *measurable* differences between the Medicaid women who received care coordination and those who did not, although full comparability is never possible. All the women on Medicaid giving birth in 1988 and 1989 had family incomes below 100% of the federal poverty level, so this was a relatively homogeneous

group in economic terms. The issue of provider selection is also salient. The maternity care providers who most actively got their patients into care coordination in the program's first years may have been among those with the strongest prenatal care programs. It is important to note, however, that the providers who implemented the service during 1988 and 1989 represented rural and urban counties and different geographic regions of the state.

The results of this study must be placed in the context of an early evaluation of a statewide system for maternity care coordination services. During 1988 and 1989, only 31% of eligible women on Medicaid received maternity care coordination. In addition, many providers were still in a start-up phase, lacking the full resources needed for the program. A goal of the Baby Love Program is to develop the service capacity necessary for providing care coordination to all eligible women. Therefore, further evaluation of maternity care coordination is needed to validate and extend the results of this study.

Low birth weight and infant mortality are serious problems in North Carolina and the United States and result from a variety of social, economic, and health care factors. Prenatal care is one important means of addressing these problems. However, there is growing evidence that, for women living in poverty, simply financing more prenatal medical visits is not enough.[7] Low-income pregnant women must also be supported in their efforts to meet basic needs such as housing, transportation, food, education, and health care. Maternity care coordination can help them meet these needs by ensuring the receipt of both medical and nonmedical services. This study suggests that maternity care coordination can be very effective in reducing low birth weight, infant mortality, and medical care costs among babies born to women in poverty. □

Acknowledgments

This paper was presented at the 118th Annual Meeting of the American Public Health Association on October 4, 1990, in New York City.

References

1. Lee K, Paneth N, Gartner LM, Pearlman M. The very low birthweight rate: principal predictor of neonatal mortality in industrialized populations. *J Pediatr.* 1980;97:759–764.
2. McCormick MC. The contribution of low birthweight to infant mortality and childhood morbidity. *N Engl J Med.* 1985;312:82–90.
3. Buescher PA. Healthier mothers and children through women's preventive health services. *NC Med J.* 1990;51:262–264.
4. Institute of Medicine. *Preventing Low Birthweight.* Washington, DC: National Academy Press; 1985.
5. Buescher PA, Smith C, Holliday JL, Levine RH. Source of prenatal care and infant birthweight: the case of a North Carolina county. *Am J Obstet Gynecol.* 1987;156:204–210.
6. Kessner DM, Singer J, Kalk CE, Schlesinger ER. Infant death: an analysis by maternal risk and health care. Washington, DC: Institute of Medicine, National Academy of Sciences; 1973.
7. Guyer B. Medicaid and prenatal care: necessary but not sufficient. *JAMA.* 1990;264:2264–2265. Editorial.

Call for Nominations: 1992 Healthtrac Prize For Improvement of Health

The Healthtrac Prize is for major achievement in health improvement, unrestricted as to field, with emphasis upon recent contributions to health in the United States, with the general criteria of achievement of the greatest good for the greatest number. It is intended for that individual who has done the most to improve health, as judged by an expert and prestigious Advisory Board.

The Prize could be awarded to a scientist, an educator, a program innovator, an activist, a public figure, a private citizen, or any other person who has made a significant contribution to improvement of the public health.

The Prize is intended to celebrate important work and bring it to public recognition so as to reward the efforts which have improved health, to increase the impact of the work, and to stimulate others to emulate the role models represented by the Prize recipient. Nominations are due April 1, 1992. The award amount is $50 000.

For further information and nomination procedures, write to Dr. James F. Fries, Healthtrac Foundation, 525 Middlefield Road, Suite 250, Menlo Park, CA 94025.

The Importance of Assessing the Fit of Logistic Regression Models: A Case Study

David W. Hosmer, PhD, Scott Taber, MS, and Stanley Lemeshow, PhD

ABSTRACT

Background. The logistic regression model is being used with increasing frequency in all areas of public health research. In the calendar year 1989, over 30% of the articles published in the *American Journal of Public Health* employed some form of logistic regression modeling. In spite of this increase, there has been no commensurate increase in the use of commonly available methods for assessing model adequacy.

Methods. We review the current status of the use of logistic regression modeling in the *American Journal of Public Health*. We present a brief overview of currently available and easily used methods for assessing the adequacy of a fitted logistic regression model.

Results. An example is used to demonstrate the methods as well as a few of the adverse consequences of failing to assess the fit of the model. One important adverse consequence illustrated in the example is the inclusion of variables in the model as a result of the influence of one subject.

Conclusions. Failure to address model adequacy may lead to misleading or incorrect inferences. Recommendations are made for the use of methods for assessing model adequacy and for future editorial policy in regard to the review of articles using logistic regression. (*Am J Public Health.* 1991;81:1630–1635)

Introduction

Modern public health research has become increasingly reliant on sophisticated statistical modeling techniques to assess the effect of new health programs, the impact of risk factors on disease, the effects of health behaviors, and a host of other public health concerns. Many such analyses involve an outcome, or dependent, variable that is dichotomous (present/absent, yes/no, live/die, etc.), and in these studies the logistic regression model has become the statistical model of choice. The reasons for this are the ease of interpretation of the estimated coefficients as "adjusted log odds ratios," the ability to estimate the probability that a particular subject will develop the outcome, and the wide availability of easily used and reliable software to perform the computations. It is now common to find, in an article using logistic regression, a table of estimated coefficients, estimated odds ratios, and associated confidence limits for the odds ratio.

The validity of inferences drawn from modern statistical modeling techniques depends on the assumptions of the statistical model being satisfied. A critical step in assessing the appropriateness of the model is to examine its fit, or how well the model describes the observed data. Very few articles in subject-matter journals make any mention of having carried out this important step in model development. Without such an analysis, the inferences drawn from the model may be misleading or even totally incorrect. Unfortunately, in some instances the results of fitting a logistic regression model have been published when the model contained serious errors that would have been detected had any attempt at goodness of fit (GOF) been performed. The logistic regression model is a powerful statistical tool, but it must be used with caution.

The goal of this paper is to present an overview of a few easily employed methods for assessing the fit of logistic regression models. An example is used to demonstrate the use of the methods and the consequences of failing to assess fit. Much of what will be presented here is described in detail in texts by Hosmer and Lemeshow[1] for logistic regression modeling and others (e.g., Kleinbaum, Kupper, and Muller[2]) for linear regression modeling. The background provided in this paper should be sufficient to allow readers to consult these texts and others for more detail.

Assessments of Goodness of Fit in the American Journal of Public Health

In order to get some idea of the frequency of use of logistic regression and methods for assessment of GOF, we reviewed the contents of the *American Journal of Public Health (AJPH)* for the calendar years 1985 through 1989. The journal's articles and brief reports were considered separately. Each published article or brief report was considered in terms of whether or not logistic regression was used and, if it was used, whether or

The authors are with the School of Public Health of the University of Massachusetts/Amherst.

Requests for reprints should be sent to David W. Hosmer, PhD, School of Public Health, University of Massachusetts/Amherst, Amherst, MA 01003.

This paper was submitted to the journal December 28, 1990, and accepted with revisions May 6, 1991.

not any assessment of GOF was performed. The results are presented in Table 1.

Overall, approximately 20% (113) of the 579 articles reviewed used logistic regression. It is interesting to note that almost 40% (45) of the papers using logistic regression appeared in 1989. Only about 5% (6) reported that any assessment of model adequacy had been performed. Logistic regression is less frequently used in brief reports (about 6%), with about the same level of attention to assessing model adequacy (about 4%).

There has been a striking increase in the use of logistic regression, but there has been no increase in the frequency of assessment of the adequacy of the model even though the techniques are now readily available in many software packages. Inferences from analyses where there has been no assessment of GOF should be viewed with some skepticism.

TABLE 1—Use of Logistic Regression (LR) and Goodness of Fit (GOF) Measures in the *American Journal of Public Health*: 1985–1989

	Articles Using LR		LR Articles Using Any Type of GOF		Brief Reports Using LR		LR Brief Reports Using Any Type of GOF	
Year	%	n	%	n	%	n	%	n
1985	16.5	85	14.3	14	2.3	86	0.0	2
1986	12.0	100	8.3	12	5.2	58	0.0	3
1987	16.0	125	0.0	20	5.3	75	0.0	4
1988	17.7	124	9.1	22	7.1	84	16.7	6
1989	31.0	145	2.2	45	10.5	76	0.0	8
Total	19.5	579	5.3	113	6.1	379	4.4	23

What Is Goodness of Fit?

Much of the lack of effort in assessing the fit of any statistical model, and logistic regression in particular, may be traced to a general misunderstanding and confusion by many users of regression methods as to what GOF is and what its role is in the modeling process.

In any statistical analysis where regression methods are employed, the goal is to obtain a mathematical model that describes the relationship between observations of the outcome or dependent variable, to be denoted as y, and a collection of independent variables, to be denoted in aggregate, without reference to their number and measurement scale, as x. A convenient way to conceptualize the model is to think of the value of the dependent variable as being composed of two parts: (1) the systematic component and (2) the error component. The systematic component is the mathematical function of the independent variables that characterizes the "average" value of y among subjects with values of the independent variables (in the model) x. The error component represents how much an individual subject's value of y differs from the "average." (We use "average" in a generic sense because, for some models, such as logistic regression, we model a transformation or function of the mean of y given x.) For example, a linear regression model for systolic blood pressure for adults might contain as independent variables height, weight, gender, and age. This portion of the model provides an estimate of the mean systolic blood pressure as a function of the variables in the model. For example, the model might give a value of 120 mm Hg for a 72-in, 160-lb, male, 40-year-old subject. The error portion for an individual measurement would be the difference between the observed value and the mean value produced by the model. The errors are often called residuals since, for many models, they are computed as the difference between the observed and predicted (systematic component) values of the dependent variable. For example, if the actual systolic blood pressure for our hypothetical subject is 110, then the value of the error would be $110 - 120 = -10$. A general formula incorporating the two components is y = systematic component + error component.

The process of "building" the regression model concentrates on the systematic component. This involves the functional relationship between the "average" of y and the content of x, the actual variables in the model. The selection of variables usually involves both biological and statistical criteria since we may include in a model variables that are of known biological importance regardless of their statistical significance. The rationale is that these variables may be needed to adjust the coefficients of other variables in the model, thereby controlling confounding. The statistical focus of model building is on the magnitude of the "slope" coefficients of the variables. Also of concern is that each variable enter the model in the correct scale. For example we should consider, for continuous independent variables, whether or not we need just x, or x^2, or $\log(x)$, etc., in the model. The result of the systematic component is a function that produces predicted or fitted values of the dependent variable. These will be denoted in general by $\hat{y}$. Attention turns to the error component after completion of the systematic component. We call the process of examining the values of the error component, both individually and collectively, assessing the GOF of the model.

Assessing GOF will usually entail two stages: computing a statistic that provides a summary measure of the errors, and examining the individual values of the errors. The summary statistic may be small yet contain individual error components that are large under the assumption of a good-fitting model.

A class of statistics called "regression diagnostics" has been proposed to examine the role of individual subjects in the model. The purpose of regression diagnostics is to aid in identifying subjects who are problematic under the current model. In this paper we consider three types of diagnostic statistics: measures of residual, leverage, and influence. A particular subject may be called to our attention because the model-based prediction of its outcome is not very close to that observed, or if it has an exceptionally large error, or "residual." Methods based on measures of residual will help identify these subjects. A subject may have a configuration of independent variables that is most unusual relative to the rest of the subjects. A measure that is used to identify these subjects is, for a variety of technical reasons, called "leverage." We are also interested in identifying those subjects who have a large effect on the systematic component as expressed by the magnitude of any of the estimated coefficients in the model. The contribution of an individual observation to the magnitude of a coefficient depends on both its residual and its leverage. These are called "influence" measures, and one form of them will be presented later.

Regardless of which diagnostic statistic is used, we must be able to ascertain which values are large. In general, histo-

grams and scatterplots are the most effective ways to view the distributions of these statistics, since what is of importance is their relative magnitude. For some types of models, such as linear regression, there are fixed cutpoints that may be used to identify an exceptionally large value for a diagnostic statistic. For other models, such as logistic regression, there are no such cutpoints, so graphical methods take on added importance. Unless a study is small, a listwise printout of the diagnostics will be too long and will not adequately present relative size.

After considering summary measures and regression diagnostics, we will have specific evidence about the fit of the proposed model. We will have an indication of how well it fits overall and how sensitive it is to individual subjects. The issues faced in dealing with problematic subjects identified using regression diagnostics will differ from study to study. Whatever approach is taken, it is important that subjects never be deleted solely on the basis of statistical criteria. In all cases, the methods used for identification and criteria for exclusion should be made clear and all subjects deleted specifically noted.

Assessing the Fit of a Logistic Regression Model

In this section we present a brief review of several basic methods for assessing the fit of a logistic regression model. For greater detail, the reader is referred to Hosmer and Lemeshow.[1] Before beginning, we need some notation to indicate the data and the fitted model.

Assume that we have n independent observations denoted by $(y_i, \boldsymbol{x}_i)$, $i = 1, \ldots, n$ where i indicates the specific subject (observation), y_i is equal to zero when the outcome is absent and is equal to one when it is present, and $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})$ contains the values of the independent variables in the model. The fitted value from the logistic regression model will be indicated by $\hat{y}_i = \hat{\pi}_i$ ($\hat{\pi}_i$ is the estimated probability that $y_i = 1$ for subject i) and is defined by the equation

$$\hat{\pi}_i = \frac{e^{\hat{g}_i}}{1 + e^{\hat{g}_i}},$$

where $\hat{g}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \ldots + \hat{\beta}_k x_{ik}$ is the sum of the estimated coefficients multiplied by the values of the independent variables for the ith subject.

We begin our discussion of assessing GOF by considering two forms of the error component for a fitted logistic regression model and their aggregate statistics. One statistic is the deviance, D, also used in model building. The deviance is given by the equation

$$D = \sum_{i=1}^{n} d_i^2,$$

where the individual components, d_i, are called the "deviance residuals." Under the assumptions in this paper, if $y_i = 1$, then

$$d_i = \sqrt{2|\ln(\hat{\pi}_i)|},$$

and if $y_i = 0$, then

$$d_i = \sqrt{2|\ln(1 - \hat{\pi}_i)|}.$$

Another summary statistic, similar to the deviance, is the Pearson chi-square statistic, which is defined as

$$\chi^2 = \sum_{i=1}^{n} r_i^2,$$

where

$$r_i = \frac{(y_i - \hat{\pi}_i)}{\sqrt{\hat{\pi}_i(1 - \hat{\pi}_i)}}$$

and the individual components, r_i, are called the "Pearson residuals." The Pearson residuals form the basis of many of the logistic regression diagnostic statistics.

The summary statistics D and χ^2 each have degrees of freedom approximately equal to $n - (p + 1)$, but unfortunately do not have an easily interpreted distribution under the assumptions of this paper, which are typical of most practical applications of logistic regression. The fact that these statistics do not follow a chi-square distribution under the hypothesis that the fitted model is, in fact, a correct model may be motivated by thinking of them as being the likelihood ratio chi-square or Pearson chi-square statistic computed from a contingency table with two rows ($y = 1$ or 0) and n columns (subjects) where the column totals are each equal to 1. The estimated expected values for the $y = 1$ row are $\hat{\pi}_i$ and, in the $y = 0$ row, $1 - \hat{\pi}_i$. Thus, any reasonable rule for large expected values would not hold and, under the assumptions of this paper, any P values that a computer program may print for these statistics should not be used. About all we can do in the way of assessing fit is to compare their values with the degrees of freedom.

One strategy that is often employed in dealing with a contingency table with small expected values is to pool biologically similar groups until the resulting categories have large enough expected values. This idea of pooling is the basis of a GOF statistic (see Hosmer and Lemeshow,[3] Lemeshow and Hosmer,[4] and Hosmer, Lemeshow, and Klar[5] for more details). This statistic is obtained by forming 10 groups of equal size containing the deciles of the fitted values. Observed and expected values are calculated by summing the estimated logistic probabilities and observed values of the outcome variable in the usual fashion. This statistic has been shown (see Hosmer and Lemeshow[3]) to follow a chi-square distribution with eight degrees of freedom when the fitted model is the correct model. This statistic is computed by a few computer software packages. The reader is referred to Hosmer and Lemeshow[1] for details on the computation of this statistic.

Any assessment of GOF should begin with the evaluation of the deviance, Pearson chi-square statistic, and, if possible, a decile-based GOF statistic. This should be followed by a careful inspection of the individual regression diagnostics to be described.

Pregibon[6] extended linear regression diagnostics to logistic regression. We present three of these diagnostics that we have found to be adequate for most studies and that may be easily computed when not provided by the computer program. As noted earlier, diagnostic statistics are of three types—leverage, residual, and influence.

The leverage of an observation was originally formulated in linear regression to measure how extreme a subject's values of the independent variables were, relative to the rest of the subjects in the study. It turned out to be a rather natural measure of the distance between an individual's values of $\boldsymbol{x}$ and the mean, $\bar{\boldsymbol{x}}$. Thus, in linear regression, the further $\boldsymbol{x}$ is from $\bar{\boldsymbol{x}}$, the greater the leverage. Leverage in logistic regression depends on weights based on the estimated logistic probability and on the distance of an observation from the mean. Define the leverage for the ith subject as $h_i = v_i l_i$, where $v_i = \hat{\pi}_i(1 - \hat{\pi}_i)$ and l_i is proportional to the distance of $\boldsymbol{x}_i$

to $\bar{x}$. The term l_i increases the further x is from $\bar{x}$. Interpretation of v_i requires some generalization for fitted logistic models. In practice, the estimated logistic probabilities, $\hat{\pi}$, either increase or decrease as a function of the independent variables. Thus, the further x is from $\bar{x}$, the larger or smaller is the estimated probability. Hence, while l_i is increasing, v_i is decreasing and approaches zero as $\hat{\pi}_i$ goes to either zero or one. The net effect is that the leverage, considered as a function of $\hat{\pi}_i$, appears quadratic in the approximate range of 0.15 to 0.85 but decreases to zero rapidly in the range of $\hat{\pi}_i < 0.15$ and $\hat{\pi}_i > 0.85$. Hence, those observations that are furthest in distance from $\hat{\pi}_i$ may have the lowest leverage. This is in direct contrast to the situation in linear regression. Maximum leverage in logistic regression tends to occur when $\hat{\pi}$ is around 0.2 or 0.8. This point will be demonstrated in the example in the next section.

Because the behavior of the leverage is closely tied to the magnitude of the estimated logistic probability, we argue that we may use the value of $\hat{\pi}$ as an indirect measure of its value in the sense that by knowing the value of $\hat{\pi}$ we know whether the leverage is expected to be either large or small. Hence, a plot of the value of leverage vs $\hat{\pi}$ assists in determining the regions where the leverage is largest.

The Pearson residuals are somewhat easier to visualize and compute than the deviance residuals, and their respective behavior is quite similar. In linear regression, we may be interested in the sign of the residual as well as its magnitude. However, in logistic regression the sign of the residual is determined by the value of y_i, as it is positive for those $y_i = 1$ and negative for those $y_i = 0$. A measure that allows us to easily determine the sign of the residual and also to more clearly see its magnitude is the amount by which χ^2 would change if the ith subject had been deleted from the model. We label this quantity Δx^2 and define it as

$$\Delta x_i^2 = rs_i^2,$$

where rs_i is the standardized Pearson residual defined by the equation

$$rs_i = \frac{(y_i - \hat{\pi}_i)}{\sqrt{\hat{\pi}_i(1 - \hat{\pi}_i)(1 - h_i)}}.$$

We can see that Δx^2 will be large when the model does not "fit" (i.e., when $y_i = 1$ and $\hat{\pi}_i$ is small, or when $y_i = 0$ and $\hat{\pi}_i$ is large). As is shown in the next section, a plot of Δx^2 vs $\hat{\pi}$ demonstrates this behavior quite clearly.

TABLE 2—Estimated Coefficients, Estimated Standard Errors, and *P* Values for the Fitted Logistic Regression Model

Variable	Estimated Coefficient	Estimated Standard Error	*P* Value
CHEM	0.526	0.249	.035
AGE	1.397	0.447	.002
SMOKE	−0.907	0.327	.006
TYPE	−1.095	0.410	.008
CHEM × TYPE	2.861	1.128	.011
Constant	−1.704	0.468	<.001

$D = 293.1$
$\chi^2 = 283.7$
$df = 279$
$\hat{C} = 22.1$, $df = 8$, $P = .005$

The effects of leverage and residual are combined to yield a diagnostic that expresses the effect of each subject on the estimated coefficients in the model. This measure was originally formulated by Cook[7] for use in linear regression and is often called "Cook distance" in regression applications. It approximates the change in the coefficient estimates due to the exclusion of the ith subject. Some programs provide this measure for each coefficient in the model, but we have found this to be unnecessarily detailed and use an overall measure to identify influential subjects. The exception is the case where there is a single coefficient in the model that is of primary importance, such as an epidemiologic study of a single risk factor. The effect on the individual coefficients may be determined by fitting a model excluding suspected influential subjects. The form of Cook distance that we prefer is given by the equation

$$\Delta\hat{\beta}_i = \Delta x_i^2 \times \frac{h_i}{(1 - h_i)}.$$

We recommend that Cook distance also be plotted vs the estimated logistic probability, as this allows one to determine the contribution of both residual and leverage to the value.

In the next section we illustrate the use of the methods for assessing the fit of a model.

A Case Study

In order to provide a realistic example, we analyze some of the data being used by a study conducted in the School of Public Health at the University of Massachusetts (X Hu, personal communication, 1989). The results of this study have not been published, so we refer to the principal variables in a generic sense. This study assessed the association between exposure to a chemical (CHEM) and the occurrence of certain symptoms (SYMP = 0 [no] or 1 [yes]) in an industrial cohort. Other important variables were determined to be whether or not the subject was a current smoker (SMOKE = 0 [no] or 1 [yes]), whether or not the subject was 40 years or older (AGE = 0 [< 40] or 1 [≥ 40]), and which of two forms of the chemical the subject was exposed to (TYPE = 0 or 1). This particular phase of the study involved 285 males, of which 93 had symptoms and 192 did not. Considerable effort went into model development to identify these as the key variables. Issues surrounding variable selection and model development have been discussed in Hosmer and Lemeshow[1] as well as in a review paper by Greenland.[8] The first model contained the main effects for each variable as well as the interaction between the chemical and its type. The results are provided in Table 2.

We can see that while D and x^2 suggest that the model might be appropriate (i.e., they are not too much larger than their degrees of freedom), the GOF statistic, $\hat{C}$, provides evidence for some lack of fit. The diagnostic plots discussed above will identify those subjects who are poorly fit and/or influential. The plots for the model in Table 2 are shown in Figures 1 through 3.

In Figure 1 we see that the points with the highest leverage values have moderate estimated logistic probabilities. As expected, the subject with the largest leverage has an estimated probability in

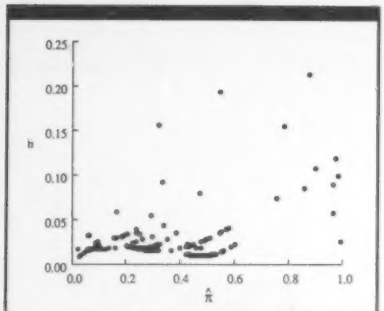


FIGURE 1—Plot of leverage vs the estimated logistic probabilities for the fitted model in Table 3.

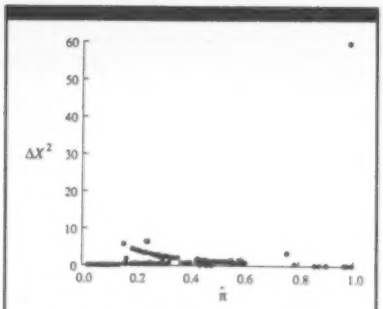


FIGURE 2—Plot of ΔX^2 vs the estimated logistic probabilities for the fitted model in Table 3.

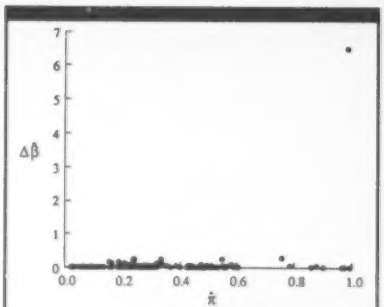


FIGURE 3—Plot of $\Delta\hat{\beta}$ vs the estimated logistic probabilities for the fitted model in Table 3.

the region around 0.8. The leverage decreases to zero as $\hat{\pi}$ approaches zero and one.

In Figure 2 the plot of the change in Pearson chi-square, Δx^2, vs $\hat{\pi}$ shows one subject who is extremely poorly fit with a value of around 60.

In Figure 3 the plot of the influence diagnostic, $\Delta\hat{\beta}$, calls our attention to this same subject. We note that this subject was not one with high leverage, indicating that the influence is primarily due to lack of fit.

At this point we examine the particular subject and see whether we can determine why he was so poorly fit. This particular individual was over 40, smoked, and had the highest level of exposure to the chemical but displayed no symptoms (i.e., $y = 0$ but $\hat{\pi}$ was quite large). If the actual data were available to us, we would probably go back to verify the exposure. In all instances, biological plausibility is an issue. Whether or not we ultimately exclude the suspected subject, it is a good idea to fit the model without the subject. The results of this fit yield a model where the estimated coefficients for the type of chemical and the type by chemical interaction are no longer significant. This suggests that a much simpler model containing only AGE, CHEM, and SMOKE could be sufficient. The results are given in Table 3.

The summary statistics indicate that the simpler model fits the data quite well when the suspected subject is deleted from the data. Plots of the diagnostic statistics support this observation. At this point, the decision as to what to do about this subject involves biological criteria that we are unable to address.

TABLE 3—Estimated Coefficients, Estimated Standard Errors, and *P* Values for the Fitted Logistic Regression Model Obtained When the Suspect Case Is Excluded

Variable	Estimated Coefficient	Estimated Standard Error	*P* Value
CHEM	2.712	0.520	<.001
AGE	1.645	0.499	.001
SMOKE	−0.914	0.344	.008
Constant	−2.808	0.524	<.001

$D = 272.6$
$x^2 = 248.2$
$df = 280$
$\hat{C} = 7.10$, $df = 8$, $P = .53$

Summary

Modern computer technology has put powerful statistical tools at the disposal of many researchers. Logistic regression is one such technique that can yield elegant results with a biologically relevant interpretation. However, the simple act of running a logistic regression package in no way guarantees that (1) the logistic model was the appropriate model to use and/or (2) the model accurately reflects the experience in the data.

Put in another context, in simple linear regression only the most naive investigator would fit a straight line to data without examining a scatterplot to determine whether a line was a reasonable model. Increasingly, researchers are using regression diagnostics to check the assumptions and assess the effect of individual cases on a multiple linear regression. There has been a growing appreciation of the fact that even one or two unusual data points can markedly affect the associations and relationships demonstrated by the vast majority of the subjects in a database. Regression texts note this point and urge that diagnostics be employed as part of any regression analysis.

While the popularity of logistic regression analysis has grown dramatically in the last decade, our review of articles published in the *AJPH* has clearly indicated that researchers may be using the fitted model without checking for its appropriateness or adequacy of fit. In 1989, 31% of the articles published in the *AJPH* used logistic regression methods, but only 2% of those used some form of GOF.

Logistic regression diagnostics provide powerful tools in the search for an accurate and useful model. Computer software packages are making these diagnostic statistics more readily available, and researchers will find that making an effort to employ them will greatly enhance their understanding of their data and enable them to build more accurate and useful statistical models.

We have shown in this paper that there are a few easily employed methods that will detect serious inadequacies in a fitted logistic regression model. The clear implication is that if a logistic model does not fit, then it does not reflect the data and thus should not be published. The editors of *AJPH* should require that authors provide some assurance that all logistic regression models have been checked for fit

and adequacy before accepting inferences based on these models. □

Acknowledgments

This paper was presented in October 1990 at the 118th Annual Meeting of the American Public Health Association, New York City.

References

1. Hosmer DW, Lemeshow S. *Applied Logistic Regression.* New York, NY: John Wiley & Sons; 1989.
2. Kleinbaum DG, Kupper LL, Muller KE. *Applied Regression Analysis and Other Multivariable Methods.* Boston: PWS-Kent; 1988.
3. Hosmer DW, Lemeshow S. A goodness-of-fit test for the multiple logistic regression model. *Commu in Stat.* 1980;A10:1043–1069.
4. Lemeshow S, Hosmer DW. A review of goodness-of-fit statistics for use in the development of logistic regression models. *Am J Epidemiol.* 1982;115:92–106.
5. Hosmer DW, Lemeshow S, Klar J. Goodness-of-fit testing for multiple logistic regression analysis when the estimated probabilities are small. *Biometrical J.* 1988;30(7):1–14.
6. Pregibon D. Logistic regression diagnostics. *Ann Statistics.* 1981;9:704–724.
7. Cook RD. Detection of influential observations in linear regression. *Technometrics.* 1977;19:15–18.
8. Greenland S. Modeling variable selection in epidemiologic analysis. *Am J Public Health.* 1989;79:340–349.

Call for Abstracts

A conference entitled the "International Child Day Care Health: Science, Prevention, and Practice," will be held at the Hotel NIKKO Atlanta, Atlanta, Georgia on June 15–17, 1992. The objective of this conference is to provide structured and informal opportunities to exchange information, skills, knowledge, and experiences related to child day care health. Presentations and discussion will focus on three major themes: (1) Child Day Care Health; (2) Meeting the Needs of Children and Care Givers; and (3) Translating Science into Practice. Submissions are encouraged for concurrent scientific sessions on infectious diseases; injuries and hazards; health promotion; children with special needs and disabilities; environmental health; developmental and psychological aspects; occupational health; impact of regulations, standards, accreditation, and training; and economics. The abstract deadline is January 15, 1992. For more information contact Lillian Glickman at Pace Enterprises, Inc. at (404) 633-8610 or FAX (404) 633-8745.

Public Health Briefs

Overcoming Denial and Increasing the Intention to Use Condoms through the Induction of Hypocrisy

Elliot Aronson, Carrie Fried, and Jeff Stone

ABSTRACT

Feelings of hypocrisy were induced in college students to increase condom use. Hypocrisy was created by making subjects mindful of their past failure to use condoms and then having them persuade others about the importance of condoms for AIDS prevention. The induction of hypocrisy decreased denial and led to greater intent to improve condom use relative to the control conditions. The implications of these findings for AIDS prevention are discussed. (*Am J Public Health.* 1991;81:1636–1638)

Introduction

Now that the acquired immunodeficiency syndrome (AIDS) epidemic is making significant inroads into the heterosexual, nonintravenous drug–using, young adult population,[1] it is vital that we convince this group to practice safer sex. The intervention of choice, thus far, has primarily been information campaigns that rely heavily on fear. Unfortunately, social psychological research has shown that fear does not always lead people toward rational behavior; it may instead trigger denial. Indeed, recent surveys have shown that, although most young adults believe that AIDS is a serious problem, they have trouble believing it is *their* problem. We have found that denial is caused by fear of AIDS coupled with a prejudice against using condoms; young people believe condoms are a nuisance that diminishes the romance and spontaneity of sexual encounters. Unless this denial can be overcome, it seems doubtful that we can convince this population to practice safer sex. If they have convinced themselves that they are not at risk, why should they change their behavior?

What kind of intervention might obviate this denial and induce young adults to realize their vulnerability and increase their resolution to practice safer sex? Imagine you are a college student and your younger brother, who is in high school, confides to you that he has become sexually active. How might you respond? You would probably urge him to use condoms. But suppose further that you were then reminded that you, yourself, do not always use condoms. How might this affect you? Recent theorizing[2] suggests that being confronted with the fact that you are not practicing what you preach induces feelings of hypocrisy, which is a form of cognitive dissonance.[3] Cognitive dissonance has been shown to produce "self persuasion," a powerful and relatively permanent form of persuasion.[4] We propose that inducing people to realize they are not practicing what they are preaching cuts off the easy route of denial and forces them to make a more realistic assessment of the risk of AIDS and, ultimately, to take adequate precautions.

Method

To test this hypothesis, it was necessary to manipulate the degree to which subjects were made aware of their own insufficient condom use and the degree to which they took an active role in preaching to others. In a two-by-two laboratory experiment, 40 female and 40 male sexually active young adults were randomly assigned to condition, counterbalanced by gender.

Upon entering the lab, all subjects were told they would be helping to develop an AIDS prevention program. Half the subjects (*high mindful*) were asked to describe fully the situations in their recent past when they failed to use condoms; half of this subgroup then went on to the preach condition. The other half of the subjects (*low mindful*) simply went directly to the preach condition without any reference to their own sexual behavior.

The authors are all with the University of California at Santa Cruz.

Requests for reprints should be sent to Elliot Aronson, Kerr Hall, University of California at Santa Cruz, Santa Cruz, CA 95064.

This paper was submitted to the journal August 30, 1990, and accepted with revisions June 24, 1991.

In the preach condition, half the high mindful and half the low mindful subjects were induced to compose a short speech (from a menu of facts) advocating condom use and to deliver it in front of a television camera. They were told this tape would be shown to high school students as part of an AIDS prevention program. The other half of the subjects (no preach) used the same menu of facts to compose a speech; these subjects rehearsed their speeches silently and were not videotaped. Thus, all subjects were exposed to the same information, but only those in the preach condition believed they were actively persuading others.

All subjects then answered questions about the frequency of their condom use in the past, as well as about their intentions to use condoms in the future. The difference between the two questions can be seen as a measure of how much an individual intends to improve his or her condom use in relation to past behaviors.

Results

The results showed support for the effectiveness of hypocrisy and are summarized in Figure 1. On the first measure, subjects in the hypocrisy condition were more likely to admit to their failure to use condoms enough in the past. This indicates that our procedure enabled subjects to overcome denial. Although a ceiling effect prevented us from documenting any possible differences on the future intentions measure, the difference between responses on the two questions indicates that the hypocrisy condition yielded a better index of improvement than any of the other conditions.

In addition to the immediate measures, we contacted subjects after 3 months and asked them about their recent condom use (Table 1). Because a sizable proportion of the subjects could not be located, statistical testing was inappropriate. Nevertheless, the obvious difference in the size of the means suggests that hypocrisy might be the most effective route to long-term behavior change.

These findings may have important implications for AIDS interventions that rely on the passive reception of communication. Specifically, when it comes to acknowledging personal risk of human immunodeficiency virus infection, simply learning about AIDS does not appear to motivate people to overcome denial and examine their risk objectively. Our information-only condition, which most closely resembles passive information, was the least effective intervention. The people within this category remained in a state of denial and were unwilling or unable to admit that AIDS was a serious threat to them personally. Additionally, learning about AIDS even when the information is directly linked to personal behavior does not appear to motivate people to examine their risk objectively, nor does simply advocating safe sex. Our data suggest that making people aware of their past high-risk behaviors must be accompanied by an engaging activity such as a public advocacy if we want them to overcome denial and adopt safer sexual behavior.

The hypocrisy technique can easily be applied to most secondary school classes on sex education or AIDS prevention. Based on our results, we suggest that lectures be supplemented by small group discussions in which each student is induced to make a public attempt to persuade others about the importance of safe sex and, subsequently, to acknowledge his or her past laxity. □

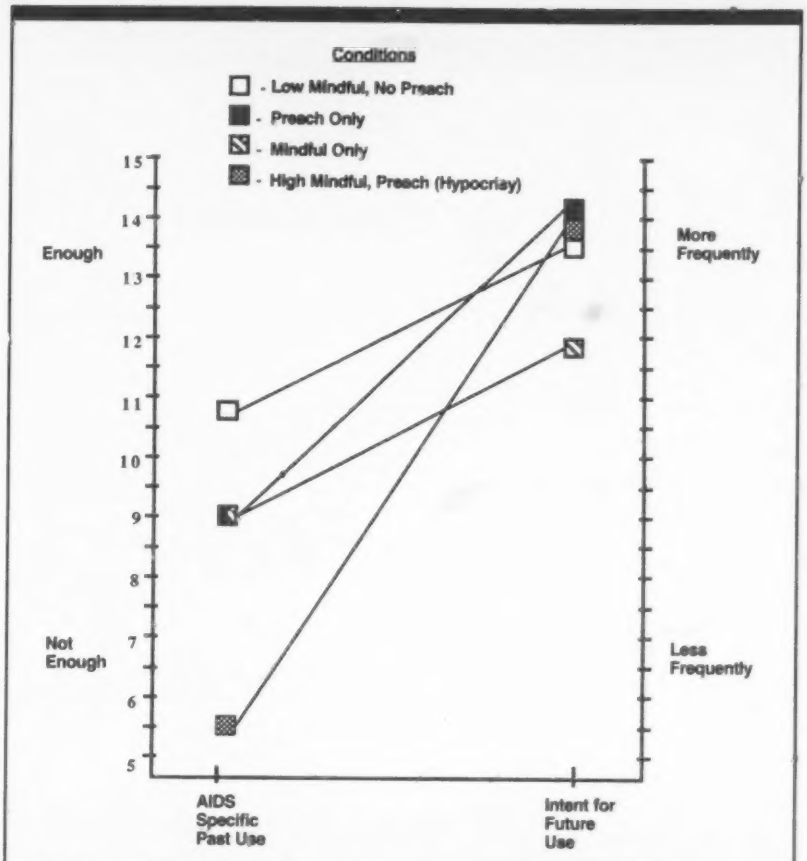


FIGURE 1—Mean responses to questions about past and future use of condoms.

TABLE 1—Sample Size and Mean Responses Indicating the Percentage of the Time Condoms Were Used Since Subjects Participated in Study

Follow-Up Interview for Self-Report at 3 Months				
	Condom Use in General		Condom Use to Prevent AIDS	
Condition	%	(n)	%	(n)
Hypocrisy	63.90	(9)	35.84	(12)
Salience only	26.25	(8)	25.00	(8)
Speech only	55.56	(9)	22.78	(9)
Information only	50.50	(6)	18.40	(10)

Acknowledgments

This research was presented at the Western Psychological Association Conference in Los Angeles, Calif, April 1990, in a paper titled "AIDS Prevention through Cognitive Dissonance: New Twist on an Old Theory."

The authors would like to thank Anthony R. Pratkanis, Judith C. Schwartz, Ruth Thi-

bodeau, and the rest of the SIRF committee for helpful comments on the design. We also thank Kyra Kissam, Jennifer Hearst, and Bruce Fraser for their help in collecting the data.

References

1. Weisse CS, Nesselhof-Kendall S, Fleck-Kandath C, Baum A. Psychosocial aspects of AIDS prevention among heterosexuals. In: Bickman L, ed. *Applied Social Psychology Annual.* Beverly Hills, Calif: Sage; 1990;10.
2. Aronson E. The return of the repressed: dissonance theory makes a comeback. *Psychol Inquiry.* In press.
3. Festinger L. *A Theory of Cognitive Dissonance.* Stanford, Calif: Stanford University Press; 1957.
4. Aronson E. Self-persuasion via self-justification: large commitments for small rewards. In: Festinger L, ed. *Retrospection on Social Psychology.* New York, NY: Oxford University Press; 1980.

Knowledge about HIV and Behavioral Risks of Foreign-Born Boston Public School Students

Ralph W. Hingson, ScD, Lee Strunin, PhD, Michael Grady, MD, Nancy Strunk, RN, MS, Robert Carr, MS, Beth Berlin, MS, and Donald E. Craven, MD

ABSTRACT

In 1990, 3049 Boston public middle and high school students were surveyed anonymously in English, Spanish, Vietnamese, Chinese, French, or Haitian Creole. Significantly fewer immigrant students, 35% of those surveyed, knew the principal modes of human immunodeficiency virus transmission. Fewer immigrants reported having sexual intercourse (31% vs 53% of nonimmigrants), but among the sexually active only 38% always used condoms, and more immigrants reported intercourse with intravenous drug users (4% vs 1% of non-immigrants). AIDS education should be taught to immigrant students in their native languages. (*Am J Public Health.* 1991; 81:1638–1641)

Introduction

Sixteen million mainland US residents age 18 and older were born elsewhere.[1] Many are Blacks or Hispanics who moved to urban areas where the incidence of AIDS has been particularly high, e.g., New York City, Miami, and Los Angeles. Little is known about their knowledge of human immunodeficiency virus-1 (HIV-1), drug use, and sexual behaviors.

Studies of adolescents particularly are needed. One fifth of the 186 895 AIDS cases nationally as of September 1991 have been in the 16 to 29-year-old age group.[2] Given the long incubation period of HIV-1, many persons in this age group probably became infected as teenagers.

This study compares middle and high school students in the Boston public schools who were born outside the US mainland with students born in the US concerning knowledge about HIV-1 transmission; beliefs about the number of adolescents who engage in risky sexual practices and drug use; and intravenous (IV) drug use, sexual intercourse, and condom use.

Methods

In May 1990, 3049 students from a random sample of Boston public schools (13/19 middle schools and 9/15 high schools) completed a self-administered questionnaire about these topics in English (n = 2704), Spanish (n = 158), Chinese (n = 45), Vietnamese (n = 50), or French or Haitian Creole (n = 92). Translation and independent back translation ensured accuracy of questionnaire wording.

We attempted to survey all 8th and 10th grade students in selected schools. In response to an informational letter about the survey sent to parents of eligible students, 75 parents (1.5%) requested that their children not participate. On the day of the survey 11% of middle school students and 23% of high school students were absent. Among students in attendance, 81% (n = 1382) in middle schools and 73% (n = 1667) in high schools completed the questionnaires, (overall response rate = 77%). Most nonresponse resulted from teachers not scheduling time to administer the survey. The gender, racial, and ethnic distributions of surveyed students closely matched those enrolled in the 8th and 10th grades in the targeted schools and in the Boston school system (Table 1). The questionnaires were anonymous and were placed by students in

Ralph W. Hingson, Lee Strunin, Beth Berlin, and Donald E. Craven are with Boston University Schools of Public Health and Medicine. Michael Grady and Nancy Strunk are with the Boston School Department. Robert Carr and Donald E. Craven are with the Boston Department of Health and Hospitals.

Requests for reprints should be sent to Ralph Hingson, ScD, Boston University School of Public Health, Social and Behavioral Sciences Section, 85 East Newton Street, Boston, MA 02118.

This paper was submitted to the journal October 30, 1990, and accepted with revisions April 2, 1991.

Editor's Note. See related Editorial by Hinman on page 1557.

blank manila envelopes. The response category, "prefer not to answer," was available for sexual and drug questions so students could not interpret the behaviors of fellow students by watching whether students answered each question. Teachers clarified student questions about the survey. Students were provided with an AIDS hotline number and instruction about AIDS was subsequently offered to all students.

Results

Students were asked to specify where they were born and how long they had lived in the United States. Thirty-five percent (n = 1057) of the students surveyed were born outside the US mainland. Half had lived in the United States less than 5 years and 21% for 1 year or less; 330 (11%) were Blacks, 309 (10%) were Hispanics, 252 (9%) were born in Asian countries, 34 (1%) were White, and the balance did not specify either their nationality or race. Twenty-one percent were born in the Caribbean or Cape Verde, 5% in Central or South America, and 8% in Asia.

The immigrant students were older ($P < .0001$): 13% compared to 5% of those born in the United States were 18 years or older. Neither the proportions in middle school or high school nor the gender distribution differed appreciably between those born in the US mainland and elsewhere.

Immigrant students were more likely to worry about getting AIDS, but less likely than US-born students to have talked to parents about AIDS and to know ways to avoid getting AIDS, where to obtain information, or where they could be tested for HIV-1 (Table 2).

Immigrant students were more likely to hold a variety of misconceptions about HIV-1 transmission. They were less likely to know that it can be transmitted by sharing needles, heterosexual intercourse, sex between two men, and sex with someone who looks healthy. They were also less likely to know ways to avoid HIV-1 transmission (Table 3). Further, they were more likely to believe that no one their age who has sexual intercourse uses condoms and that most or all people their age are injecting drugs (Table 4).

Immigrant students were less likely to report ever having had sexual intercourse. However, among those immigrants who were sexually active, a higher percentage reported having had sex with an IV drug user. Regardless of whether they were born in the United States or not, 62% of sexually active adolescents did not consistently use condoms and 2% injected illegal drugs (Table 4).

Hispanics, Blacks, Asians, and Whites born outside the United States were less knowledgeable about HIV transmission than students in these

TABLE 1—Racial and Ethnic Background of 8th- and 10th-Grade Students, Boston Public Schools in 1990 AIDS Survey, Surveyed Schools and System-wide

Background	Students Surveyed (N = 3049)	8th and 10th Graders in Survey Schools (n = 4842)	8th and 10th Graders in Boston Public Schools (n = 8182)
White	20	19	22
Black	44	49	49
Hispanic	20	22	19
Asian	10	9	10
Other (unspecific)	6	<1	<1

Note. Numbers are percentages of total.

TABLE 2—Differences in Responses to Selected Survey Questions among Students Born in US Mainland and Immigrant Students in Boston Public Schools

Question	US-Born (n = 1963)	Immigrant (n = 1057)	Odds Ratio (95% CI)
Worry a great deal about			
Getting AIDS	31	41	1.35 (1.11–1.65)
Discussed AIDS with parents	60	51	0.69 (0.59–0.80)
Where to get AIDS information	72	50	0.40 (0.34–0.46)
Ways to avoid getting AIDS	92	71	0.21 (0.17–0.26)
Where to get tested for HIV-1	56	38	0.50 (0.42–0.58)
Erroneous Beliefs			
A person would have to be sick to infect others	25	33	1.47 (1.24–1.73)
There is a cure for AIDS	2	7	3.25 (2.23–4.75)
Only gay men get AIDS	3	8	2.42 (1.72–3.41)

Note. Numbers are percentages of total answering "yes"; place of birth unspecified, n = 29; CI = confidence interval.

TABLE 3—Knowledge about HIV Transmission among Students Born in US Mainland and Immigrant Students in the Boston Public Schools

Question	US Born (n = 1963)[a]	Immigrant (n = 1057)[a]	Odds Ratio (95% CI)
Can get AIDS from			
Insect bites	17	24	1.55 (1.28–1.87)
Shaking hands	1	6	4.42 (2.79–7.01)
Being in same room with a person with AIDS	1	10	8.47 (5.38–13.33)
Sex between a man and a woman	94	85	0.32 (0.25–0.42)
Sex between two men	84	76	0.59 (0.49–.071)
Sex with someone who looks healthy	59	43	0.53 (0.45–0.62)
Sharing needles when injecting drugs	97	87	0.22 (0.17–0.31)
Can avoid AIDS by			
Sexual abstinence	73	56	0.37 (0.32–0.44)
Using condoms	87	70	0.35 (0.29–0.43)
Having sex with 1 person who is not infected	70	63	0.73 (0.62–0.85)
Not having sex with an IV drug user	73	64	0.66 (0.56–0.78)

Note. Place of birth unspecified, n = 29; CI = confidence interval.
[a]Percent of subsample responding "yes."

TABLE 4—Perceptions of Peer Behavior and Personal Sexual and Drug-Use Practices of Students Born in US Mainland and Immigrant Students in the Boston Public Schools

Question	US Born (n = 1983)[a]	Immigrant (n = 1057)[a]	Odds Ratio (95% CI)
Perceptions of peer behavior			
How many people your age are having sex?	64[b]	52[b]	0.61 (0.52–0.71)
How many people your age are using condoms?	6[c]	14[c]	2.29 (1.84–3.19)
How many people your age are injecting drugs?	12[b]	28[b]	2.76 (2.21–3.45)
Personal risk behaviors			
Ever had sex	53	31	0.38 (0.30–0.43)
Had sex in the past month with an IV drug user	1	4	2.86 (1.29–6.34)
Use condom because of AIDS risk	66	69	0.83 (0.65–1.14)
Use condoms past month	38[d]	38[d]	0.62 (0.45–0.86)
Use IV drugs past month	2	2	0.96 (0.54–1.71)
Ever shared needles	1	1	0.85 (0.36–2.01)

Note. Place of birth unspecified, n = 29; CI = confidence interval.
[a]Percentage of subsample responding "yes," except as noted.
[b]Percent responding "almost all or most."
[c]Percent responding "none."
[d]Percent responding "all the time."

groups born in the United States (Table 5). Students who had immigrated to the US mainland in the previous year were less knowledgeable about AIDS and more likely to believe that their peers inject drugs and do not use condoms than those who had lived here 5 or more years. These newer immigrants were also more likely to have had sex with an IV drug user and to engage in sexual intercourse without using condoms (Table 6).

Conclusion

While most adolescents nationally[3] and in Massachusetts[4,5] know the principal modes of HIV transmission, those in racial and ethnic minorities are less knowledgeable,[6–9] particularly students from other countries and Puerto Rico now residing in the United States. It should be noted that our study reflects the perceptions and practices of immigrant students in the Boston public schools. Knowledge and behaviors may differ in other cities. The use of self-administered surveys often yields lower response rates and more missing data than in-person interviews. Also, absent students may have been less knowledgeable and more likely to engage in risky behaviors.

Despite these limitations, our findings underscore the importance of offering AIDS education in multiple languages in school systems that serve large immigrant populations. Students born outside the United States may bring varying notions about disease causation, prevention, and treatment, and educators should explore those beliefs in a nonthreatening fashion.[10]

Immigrant students may walk a tightrope between acculturation pressures to assume behaviors they perceive as commonplace among peers born in the United States and parental pressure to retain their cultural values and traditions. Students who immigrated to the United States in the previous year were more likely than those living here longer to believe their peers engage in unprotected sex and use IV drugs. Recent immigrants who want to "fit in" with perceived US social norms may be particularly vulnerable to pressure from other teens or adults to engage in unprotected sex or to use illicit drugs.

These students should be helped to understand that unprotected sex and illicit drug use are not as common as they may suspect. Peer-led instruction from students from their own countries who have lived in the United States for several years might be especially helpful, as would training in skills to resist peer pressure.

TABLE 5—Knowledge about HIV-1 Transmission among Boston Public School Students According to Race and Nation of Birth

	Black		Hispanic		Asian		White		Other	
Question	US (n = 974)	I (n = 330)	US (n = 272)	I (n = 309)	US (n = 52)	I (n = 252)	US (n = 568)	I (n = 34)	US (n = 67)	I (n = 113)
Know										
Ways to prevent getting AIDS	92	78	89	84	92	49	94	77	96	66
Where to get information about AIDS	76	57	68	61	67	30	65	44	80	47
Can get AIDS from										
Sex between men and women	93	87	95	85	96	79	96	84	97	90
Sex between two men	82	82	79	80	82	69	90	88	88	65
Sex with someone who appears healthy	54	46	48	44	73	39	71	50	73	39
Sharing needles	96	85	97	91	100	83	96	85	97	87
Toilet Seats	5	11	7	14	6	13	6	15	2	18
Being in the same room with a person who has AIDS	1	8	2	6	0	15	1	6	0	14

Note. US = survey students born in the mainland United States; I = immigrant students; numbers are percentages of totals answering "yes."

The need for such education is magnified by findings that, like US-born students, most sexually active immigrant students do not use condoms and, although this sexually active group is small in absolute numbers, a higher proportion reported having had sex with an IV drug user compared with US-born students. □

TABLE 6—Selected Beliefs about AIDS and Sexual and Drug Use Practices of Students Who Lived in the US Mainland ≤1 Year Compared to Students Who Have Lived in the United States >5 Years

Question	Lived in US ≤1 yr (n = 195)[a]	Lived in US >5 yr (n = 369)[a]	Odds Ratio (95% CI)
Worry a great deal about getting AIDS	45	35	1.54 (1.08–2.19)
Can get AIDS from being in the same room with an infected person	20	4	6.69 (3.45–12.94)
Abstaining from sex can reduce chances of getting AIDS	56	68	0.59 (0.41–0.86)
Using condoms can reduce chances of getting AIDS	60	83	0.31 (0.21–0.46)
What proportion of people your age use condoms[b]	19	9	2.41 (1.40–4.14)
Almost all/most people your age inject drugs	32	21	1.71 (1.13–2.58)
Use condoms because of AIDS risk	50	72	0.39 (0.19–0.82)
Always used condoms past year	32	49	0.49 (0.23–1.05)
Had sex with an IV drug user in the last year	13	1	10.30 (1.91–55.47)

Note. CI = confidence interval.
[a]Percentage of subsample responding "yes," except as noted.
[b]Percentage responding "none."

Acknowledgments

This project was supported by cooperative agreement contract No. U63/CCU102792 to the Boston Public Schools from the Division of Adolescent School Health, Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control, Atlanta, Ga.

This paper was presented, in part, at the 118th Annual Meeting of the American Public Health Association, in New York City, October 1, 1990.

References

1. Hendershot G. Health of the foreign born population, United States 1985–1986. *Advance Data* no. 157. National Center for Health Statistics; 1988.
2. *AIDS Newsletter Monthly,* Publication of the Massachusetts Dept. of Public Health 7;3, September 1991;7:3.
3. Hardy A. AIDS knowledge and attitudes for April–June 1989. *Advance Data* no. 179. National Center for Health Statistics; 1989.
4. Hingson R, Strunin L, Berlin B. AIDS transmission: changes in knowledge and behaviors among adolescents: Massachusetts statewide surveys 1986–1988. *Pediatrics.* 1990;85:24–29.
5. Hingson R, Strunin L, Berlin B, Heeren T. Beliefs about AIDS, use of alcohol and drugs, and unprotected sex among Massachusetts adolescents. *Am J Public Health.* 1990;80:295–300.
6. Dawson D, Hardy A. AIDS knowledge and attitudes of Hispanic Americans. *Advance Data* no. 166. National Center for Health Statistics; 1989.
7. Dawson D, Hardy A. AIDS knowledge and attitudes of Black Americans. *Advance Data* no. 165. National Center for Health Statistics; 1989.
8. DiClemente RJ, Boyer L, Morales E. Minority and AIDS: knowledge, attitudes, and misconceptions among Black and Latino adolescents. *Am J Public Health.* 1988;78:55–57.
9. Goodman E, Cobrall A. Acquired immunodeficiency syndrome and adolescents: knowledge, attitudes, beliefs, and behaviors in a New York City adolescent minority population. *Pediatrics.* 1989;84:36–42.
10. Strunin L. Adolescents perceptions of risk for HIV infection: implications for future research. *Soc Sci Med.* 1991;2:221–228.

HIV Infection and Risk Behaviors among Intravenous Drug Users in Low Seroprevalence Areas in the Midwest

Harvey A. Siegal, PhD, Robert G. Carlson, PhD, Russel Falck, BS, Ling Li, MBA, Mary Ann Forney, PhD, Richard C. Rapp, MSW, Kathy Baumgartner, MS, William Myers, MS, and Morton Nelson, MD, MPH

ABSTRACT

We studied behavioral factors that place intravenous drug users at risk for the acquisition and transmission of the human immunodeficiency virus (HIV) in a sample of 855 individuals not in drug treatment, living in central and southwestern Ohio. The HIV seropositivity rate for the sample was 1.5%. Three factors were significantly related to HIV infection: homeless shelter residence (odds ratio [OR] = 7.7, 95% confidence interval (CI) = 3.0–20.0), travel to northeastern HIV hyperendemic areas (OR = 5.2, 95% CI = 1.8–15.4), and recent male homosexual or bisexual behavior (OR = 11.2, 95% CI = 2.9–43.9). (*Am J Public Health.* 1991;81:1642–1644)

Introduction

Intravenous drug users (IVDUs) are at high risk for contracting the human immunodeficiency virus (HIV).[1] In the United States, epidemiologic studies have focused primarily on assessing the risk behaviors of IVDUs in high (>15%) HIV seroprevalence areas. Risk factors associated significantly with high HIV seropositivity include the following: minority status, male homosexual or bisexual behavior, a history of sexually transmitted diseases (STDs), frequency of injection, cocaine injection, drug use in shooting galleries, and sharing of drug injection equipment.[2–5] While studies of IVDUs in low seroprevalence areas have identified similar risk behaviors, these have generally not been associated significantly with positive serostatus.[6–9]

This paper first describes the prevalence of HIV risk behaviors reported by 855 active IVDUs in a low seroprevalence area of the Midwest. It then examines the relationships among sociodemographic variables, residence and mobility patterns, drug-use patterns, sexual behaviors, and HIV infection.

Methods

From March 1989 to July 1990, 855 IVDUs were recruited from the streets in the metropolitan areas of Columbus and Dayton, Ohio, by indigenous outreach workers. To be eligible for the study, subjects had to have injected some drug at least once in the previous 6 months, be at least 18 years of age, and not have been in drug treatment in the 30 days prior to data collection.

Data were collected using the AIDS Initial Assessment Instrument (AIA), developed by the National Institute on Drug Abuse for its AIDS outreach and prevention projects. All 855 IVDUs voluntarily agreed to have a sample of their blood screened for HIV antibodies. Serum samples were initially screened using the Abbott enzyme-linked immunosorbent assay. Confirmation tests consisted of Western blotting and were considered positive when antibodies to two of three bands were detected: gp 160/120, gp41, p24.

Based on the results of descriptive statistics, Mantel-Haenszel odds ratios (OR) with 95% confidence intervals (CI) were calculated to compare seropositive and seronegative IVDUs on selected variables. Chi-square tests corrected for continuity were employed to assess significant associations.

Results

Males constituted 75% of the sample. Seventy-eight percent of the sample were Black, 21% were White, and 1% belonged to other racial groups. The median age was 36.

Drug and Needle Use Behaviors

Striking in this sample are the high rates of cocaine use: 79% had injected cocaine in the previous 6 months while 68% reported smoking crack cocaine occasionally, weekly, or daily. Sixty-one percent had injected heroin, and 43% had injected heroin mixed with cocaine (speedball).

About half (49%) reported sharing needles or syringes within the last 6 months. Only 31% reported injecting drugs in a shooting gallery, and almost one quarter (23%) of the IVDUs claimed that they always used a new needle and syringe.

The research for this paper was conducted at Wright State University, School of Medicine, Substance Abuse Intervention Program, Dayton, Ohio; Columbus Health Department, Columbus, Ohio; and Montgomery County Combined Health District, Dayton, Ohio.

Requests for reprints should be sent to Harvey A. Siegal, PhD, Professor and Director, Wright State University, School of Medicine, Substance Abuse Intervention Programs, PO Box 927, Dayton, OH 45401.

This paper was submitted to the journal November 15, 1990, and accepted with revisions May 6, 1991.

TABLE 1--Serostatus and Risk Factors

Factor	n	Serostatus: % Positive	Serostatus: % Negative	χ_0^{2a}	*P* Values	Odds Ratio	95% Confidence Interval
Sex							
Male	639	1.6	98.4	0	1	1.1	.31–4.1
Female	216	1.4	98.6				
Ethnicity							
Black	664	1.5	98.5	0	1	.95	.26–3.5
White	189	1.6	98.4				
Daily injection							
Yes	297	2.4	97.6	1.36	.24	2.2	.76–6.5
Daily cocaine injection							
Yes	150	3.3	96.7	2.66	.10	3.0	1.0–8.8
Crack use							
Yes	584	1.4	98.6	.07	.79	.72	.24–2.2
Use of shooting gallery							
Yes	330	1.8	98.2	.07	.79	1.4	.46–4.1
Always used new needles							
No	655	1.8	98.2	1.04	.31	3.7	.55–25.1
Ever shared or rented needles							
Yes	421	2.1	97.9	1.38	.24	2.4	.74–7.4
Homeless shelter residence							
Yes	90	6.7	93.3	14.12	.000	7.7	3.0–20.0
Epicenter travel[b]							
Yes	70	5.7	94.3	6.17	.01	5.2	1.8–15.4
History of STDs							
Yes	336	2.4	97.6	1.84	.18	2.5	.84–7.4
Number of sex partners							
Multiple	452	1.6	98.4	.82	.37	2.6	.57–12.0
Condom use							
Yes	216	1.9	98.1	.61	.44	2.1	.59–7.8
Homo/bisexual males[c]							
Yes	22	9.1	90.9	6.17	.01	11.2	2.88–43.8

[a]Chi-square test corrected for continuity.
[b]For this study northeastern HIV epicenters include New York, NY, Newark, NJ, and Jersey City, NJ.
[c]The AIDS Initial Assessment instrument is designed to assess risk behaviors in the 6 months preceding the interview. If a respondent reports no sexual activity, there is no mechanism to determine sexual orientation. Of the 10 seropositive males, sexual orientation is known for 7 participants: 5 report heterosexual behavior and 2 report homosexual behavior. Given these small numbers, results must be interpreted cautiously.

Sexual Practices

The majority of IVDUs (93%) reported having sexual relations within the previous 6 months; 58% indicated having multiple partners. Almost 40% of the IVDUs reported a history of at least one STD. Of all those IVDUs sexually active in the previous 6 months, 96% reported sexual behaviors that we categorized as heterosexual, 2% as bisexual male (3% female), and 2% as homosexual male (4% female).

HIV Seropositivity and Risk Behavior

The seropositivity rate for the total sample was low—1.5% (13 of 855 subjects). A greater percentage of seropositive subjects injected drugs on a daily basis compared to seronegative participants (54% vs 34%, respectively), reported a higher daily cocaine injection rate (38% vs 17%), and were more likely to live in a homeless shelter (46% vs 10%), to travel to HIV epicenters (31% vs 11%), to have multiple sex partners (78% vs 57%), to have a history of STDs (62% vs 39%), to have ever shared needles or syringes (69% vs 48%), and to use new needles less frequently (7% vs 24%).

Three variables were associated significantly ($P < .05$) with seropositivity (Table 1). Those persons living in a shelter were seven times more likely to be infected (OR = 7.7, 95% CI = 3.0–20.0), while those who had traveled to a northeastern HIV hyperendemic area within the previous 2 years were five times more likely to be seropositive (OR = 5.2, CI 1.8–15.4). Finally, homosexual/bisexual males who were sexually active in the 6 months prior to the interview were 11 times more likely to be HIV positive (OR = 11.2, CI = 2.9–43.9).

Discussion

Our results are generally similar to previous research in low seroprevalence areas.[6–9] The widespread practice of AIDS-related risk behaviors in this midwestern locale provides the potential for higher infection rates in the future. At present, however, variables typically associated with HIV infection in high seroprevalence areas, such as race and ethnicity, use of shooting galleries, and frequency of injection had little predictive power, although many results are in the expected direction. In contrast with previous studies,[10–11] Whites and Blacks were about equally likely to be seropositive.

While our analyses are based on a small number of seropositive cases and should be interpreted cautiously, several new findings emerged. The results encourage some speculation about the movement of HIV to low seroprevalence areas as well as the key mechanisms for local dispersion. The significant relationship between travel to northeastern hyperendemic areas and HIV seropositivity is a new finding among IVDUs in low seroprevalence areas.

This study seems to support the findings of Battjes and colleagues[6] that homosexual/bisexual male IVDUs form a potential route for transmission of the HIV from non-drug-using gay males to heterosexual IVDUs in low seroprevalence areas. Our data emphasize the importance of travel as an additional mechanism for importing the virus into low seroprevalence areas. If the diffusion factors of mobility and homosexual/bisexual behavior are interrelated, future increases in the prevalence of HIV in currently low seroprevalence regions may be expected, given the high levels of risk behaviors. Such scenarios, tentative as they are, emphasize the critical need for preventive interventions in low seroprevalence areas.

Our results also suggest that homeless shelter residence is a potentially high-risk social setting that may serve as an additional indirect mechanism for the dispersion of HIV. While a high seroprevalence rate has been reported among male IVDUs residing in homeless shelters in New York City,[12] the association between shelter residence and seropositivity is a new finding for low seroprevalence areas and might contribute to understanding the natural history of HIV. Shelters bring together many of society's most vulnerable people, allowing intense social interaction. Given the current crack cocaine epidemic with its associated high-risk sexual behavior and disruption of social networks,[13–14] these interrelated processes are cause for concern.

Future research should examine IVDU mobility, homosexual/bisexual behavior, and social networks particularly as these interface with homeless shelter residence patterns. Focus should be directed toward understanding these factors within and between low and high seroprevalence regions. □

Acknowledgments

This research was supported by grant DA05757 from the National Institute on Drug Abuse.

The authors wish to acknowledge the contributions of Ms. Ruth Frankenfield, RN, MSN, City of Columbus site director, Ms. Diana Alexander, Dayton site director, and members of their staffs.

References

1. Berkelman RL, Heyward WL, Stehr-Green JK, Curran JW. Epidemiology of human immunodeficiency virus infection and acquired immunodeficiency syndrome. *Am J Med.* 1989;86:761–770.
2. Marmor M, Des Jarlais DC, Cohen H, et al. Risk factors for infection with human immunodeficiency virus among intravenous drug abusers in New York City. *AIDS.* 1987;1:39–44.
3. Chaisson RE, Bacchetti P, Osmond D, Brodie B, Sande MA, Moss AR. Cocaine use and HIV infection in intravenous drug users in San Francisco. *JAMA.* 1989;261:561–565.
4. Schoenbaum EE, Hartel D, Selwyn PA, et al. Risk factors for human immunodeficiency virus infection in intravenous drug users. *N Engl J Med.* 1989;321:874–879.
5. McCoy CB, Khoury E. Drug use and the risk of AIDS. *Am Behav Sci.* 1990;33:419–431.
6. Battjes RJ, Pickens RW, Amsel Z. Introduction of HIV infection among intravenous drug abusers in low prevalence areas. *J Acquir Immune Defic Syndr.* 1989;2:533–539.
7. Chandrasekar PH, Molinari JA, Kruse JA. Risk factors for human immunodeficiency virus infection among parenteral drug abusers in a low prevalence area. *South Med J.* 1990;83:996–1001.
8. Wendt D, Sadowski L, Markowitz N, Saravolatz. Prevalence of serum antibody to human immunodeficiency virus among hospitalized intravenous drug abusers in a low-risk geographic area. *J Infect Dis.* 1987;155:151–152. (Letter).
9. Williams ML. HIV seroprevalence among male IVDUs in Houston, Texas. *Am J Public Health.* 1990;80:1507–1509.
10. Selik RM, Castro KG, Pappaioanou M. Racial/ethnic differences in risk of AIDS in the United States. *Am J Public Health.* 1988;78:1539–1545.
11. Lange R, Snyder FR, Lozovsky D, Kaistha V, Kaczaniuk MA, Jaffe JH. Geographic distribution of human immunodeficiency virus markers in parenteral drug abusers. *Am J Public Health.* 1988;78:443–446.
12. Torres RA, Sridhar M, Altholz J, Brickner PW. Human immunodeficiency virus infection among homeless men in a New York City shelter. *Arch Intern Med.* 1990;150:2030–2036.
13. Inciardi JA. Trading sex for crack among juvenile drug users: a research note. *Cont Drug P.* Contemporary Drug Problems 1989;16:689–700.
14. Carlson RG, Siegal HA. The crack life: an ethnographic overview of crack use and sexual behavior among African-Americans in a midwest metropolitan city. *J Psychoactive Drugs.* 1991;23:11–20.

Frequency and Thoroughness of STD/HIV Risk Assessment by Physicians in a High-Risk Metropolitan Area

Bradley O. Boekeloo, PhD, ScM, Eric S. Marx, MTS, Alex H. Kral, BA, Steven C. Coughlin, PhD, Marjorie Bowman, MD, MPA, and David L. Rabin, MD, MPH

ABSTRACT

The US Preventive Services Task Force recommends that all primary care physicians assess the sexually transmitted disease/human immunodeficiency virus (STD/HIV) risk of all adolescent and adult patients. To determine whether factors amenable to change through continuing medical education are associated with frequent and thorough STD/HIV risk assessment, a telephone survey of primary care physicians in the Washington, DC metropolitan area was conducted (n = 961). Thirty-seven percent of physicians reported regularly asking new adult patients about their sexual practices; 60% asked new adolescent patients. STD/HIV risk questioning was associated with physicians' confidence in their ability to help prevent HIV, comfort with discussing patients' sexual practices, and perception of a large STD/HIV problem in their practice. These findings suggest that continuing medical education should target improvement in physicians' sexual practice questioning skills. (*Am J Public Health.* 1991;81:1645–1648)

Introduction

Because of recent increases in the number and incidence of sexually transmitted diseases (STDs) and the lethality of human immunodeficiency virus (HIV) infection, improved physician STD/HIV prevention practice is urgent.[1,2] The US Preventive Services Task Force recommends that all primary care clinicians obtain a complete sexual practice and drug-use history from all adolescent and adult patients.[1] Primary care physicians report however, that they infrequently assess the sexual practices of their patients.[3]

Recent studies indicate that discomfort with homosexuality is a barrier to HIV-related care.[3–5] Non-internist, older, male, solo-practice physicians and physicians having no AIDS patients are all reported to have less HIV-related competence.[3–5] Although many barriers to STD/HIV risk assessment cannot be overcome through continuing medical education (CME) (i.e., physician's age or type of practice), CME can potentially decrease barriers caused by lack of knowledge, predisposing attitudes, and skills. In this study, current levels of self-reported primary care physician STD/HIV risk questioning are measured and educationally mutable factors associated with this questioning are identified.

Methods

Baseline data for a demonstration project of physician office-based STD/HIV prevention programs using simulated patients[6] were obtained through telephone interview. The sampling frame was the 1988 American Medical Association (AMA) list of office-based practicing physicians whose mailing addresses were in the Washington DC metropolitan statistical area.

Physicians were eligible for participation in the study if (1) they identified themselves at the time of the interview as being in one of four designated primary care specialties (internal medicine, family practice, general practice, and obstetrics/gynecology), and (2) they delivered at least 1 hour of primary care each week. The overall survey response rate was 60% (N = 961).

All telephone survey questions were closed-ended, with trained interviewers recording responses on precoded forms. In data analysis, variable associations were measured using Pearson product-moment correlation coefficients and forward stepwise multiple linear regression (PROC REG).[7]

Independent Variables

Study physicians were asked about their personal and practice characteristics. They were also asked to rate the magnitude of the STD/HIV problem in their practice. A battery of seven 5-point Likert scale questions (available from the authors) was used to measure physicians' confidence in their ability to assess risk and successfully counsel patients in the prevention of HIV infection (Cronbach alpha = 0.68, range = 7–35, mean = 25.0, SD = 5.00, median = 25). Physicians were also asked about their comfort level when discussing sexual practices with five different types of patients: single women, homosexual women, long-time married patients, homosexual men, and single men. The cronbach A among these items was 0.78. These 3-point item scores were

All authors are with Georgetown University School of Medicine, Department of Community and Family Medicine, Washington, DC, except for Marjorie Bowman, who is with Bowman Gray School of Medicine, Department of Family and Community Medicine, Winston-Salem, NC.

Requests for reprints should be sent to Bradley O. Boekeloo, PhD, ScM, Georgetown University School of Medicine, Department of Community and Family Medicine, 3750 Reservoir Road, NW, Kober-Cogan Room 204, Washington, DC 20007.

This paper was submitted to the journal November 15, 1990, and accepted with revisions May 6, 1991.

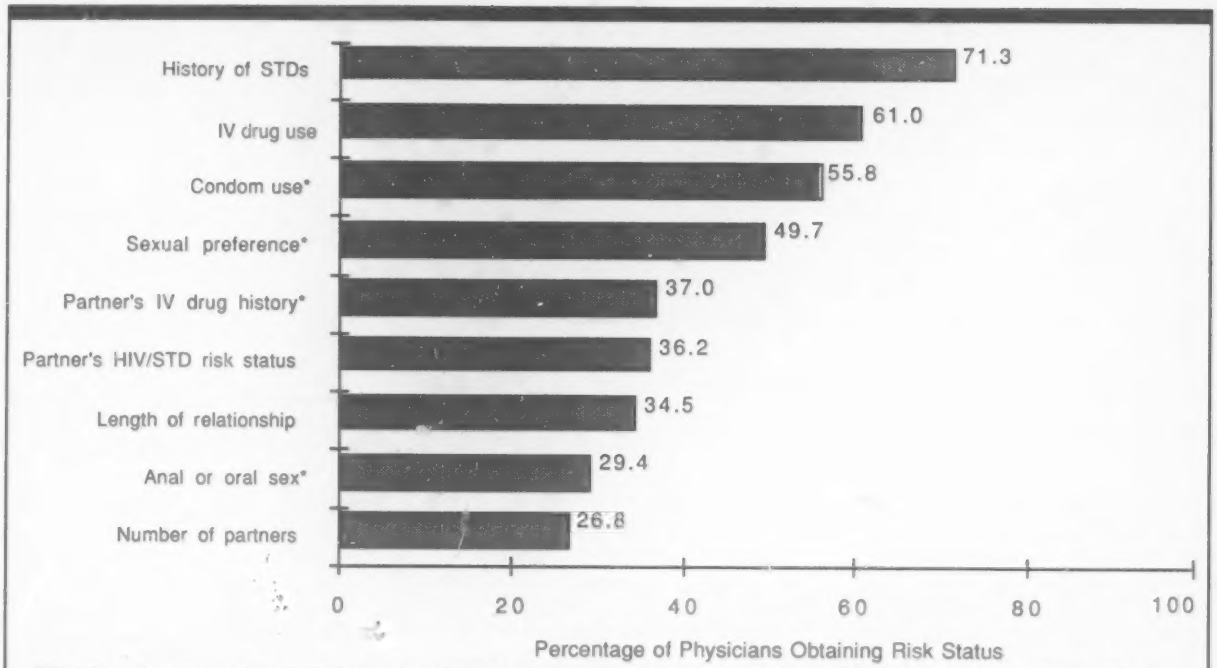


FIGURE 1—Percentage of Physicians (n = 961) Reporting They Always/Often Obtain HIV Risk Status of New Patients, by HIV Risk Factor.

*This question was asked only about new patients whom the physician knew had more than one sexual partner in the last ten years.

totaled and divided by 5 to create an overall Discomfort Scale. For obstetricians/gynecologists the scale excluded questions regarding male patients and total scores were divided by 3. Because the Discomfort Scale scores were skewed, they were categorized into levels of high, moderate, and low discomfort.

Dependent Variables

Physician STD/HIV risk questioning, the major outcome variable, was measured by asking physicians how often they ask new patients during initial general health examinations about nine sexual STD/HIV risk factors on a 5-point scale (Never-Always) (Figure 1). A summative scale of these items formed the STD/HIV Risk Questioning Scale (SRQ) (Cronbach alpha = 0.85, range = 9–45, mean = 28.46, SD = 7.75, median = 28). Frequency of asking 12 different types of patients about their sexual practices or changes in sexual practices was also obtained (Figure 2) on a 5-point scale (Never-Always).

Results

Characteristics of the Study Population

Study physicians were predominantly White (72%), male (80%), board certified in their primary care specialty (74%), and graduates of US medical schools (76%). Their average age was 47 years (SD = 10.61) and 48% were in solo practice. Sixty-five percent had diagnosed an HIV-positive patient. Information from the AMA indicated small but significant differences between study participants and nonparticipants on two of five variables: gender and practice organization (i.e., solo, partnership, group).

Physician STD/HIV Risk Questioning

During initial general health examinations of new patients, the majority of physicians reported regularly ("always" or "often") asking about history of STDs (71%) (Figure 1). Fewer physicians reported regularly asking about other STD/HIV risk factors such as condom use (56%), sexual preference (50%), anal or oral sex (29%), or number of partners (27%).

The percentage of physicians who reported regularly questioning patients about sexual practices varied widely by patient type (Figure 2). Thirty-seven percent of primary care physicians reported regularly asking their new adult patients about their sexual practices; 60% asked their new adolescent patients. The majority of physicians reported regularly asking about sexual practices of patients in traditional HIV risk groups—homosexual men (87%) and IV drug users (81%)—and patients who asked them about AIDS (92%).

Associations with STD/HIV Risk Questioning

SRQ scores were positively associated with several "immutable" characteristics of physicians and their practices in bivariate and multivariate analyses (Table 1) including being a foreign medical school graduate ($r = .12$, $p \leq .0002$), being female ($r = .12$, $p \leq .0003$), and having had four or more HIV-positive patients ($r = .10$, $p \leq .002$).

SRQ scores were also positively associated with "potentially mutable" characteristics of physicians in bivariate and multivariate analyses (Table 1) including perceiving STD/HIV infection to be a big problem in one's practice ($r = .23$, $p \leq .0001$), being comfortable when discussing sexual practices with patients ($r = .24$, $p \leq .0001$), and having greater STD/HIV prevention confidence ($r = .29$, $p \leq .0001$).

Discussion

This study indicates that in one of the highest STD/HIV incidence areas in the country, metropolitan Washington, DC, primary care physician STD/HIV risk assessment rates are slightly higher than those recently reported in other high-risk urban areas for Los Angeles,[8,9] but are well

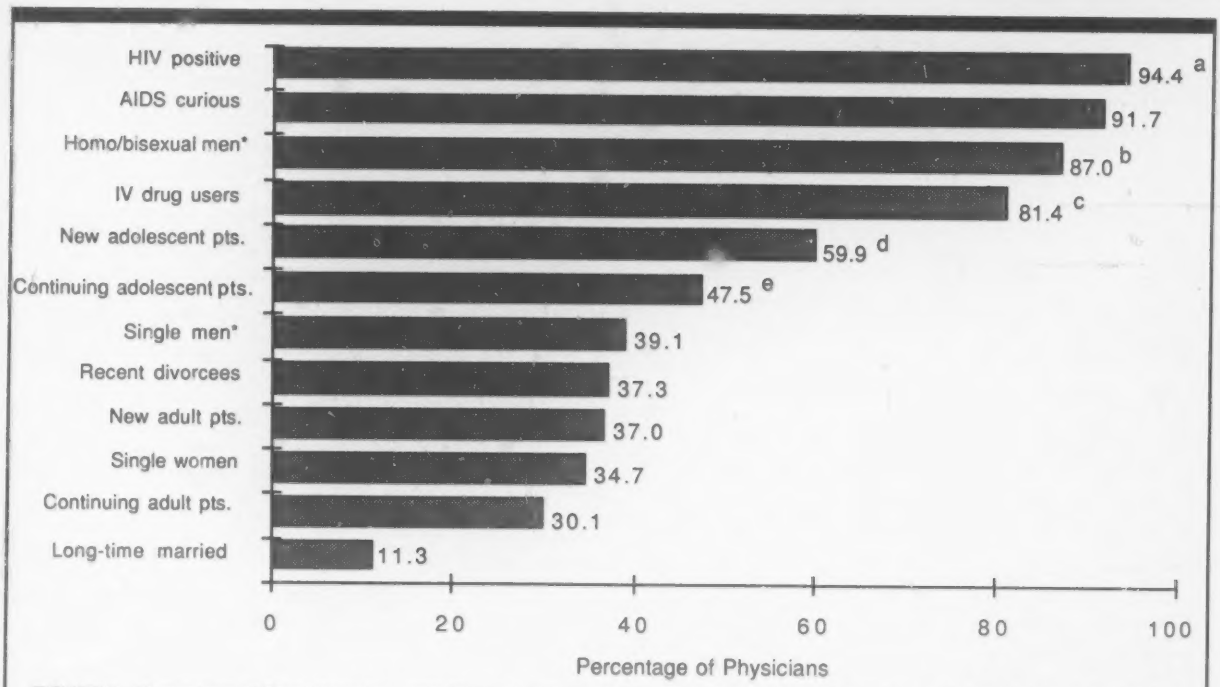


FIGURE 2—Percentage of Physicians (n = 961) Reporting They Always/Often Ask Patients About Their Sexual Practices, by Patient Category.

*OB/GYNs not asked, n = 693
a. 204 physicians stated they do not have this type of patient.
b. 47 physicians stated they do not have this type of patient.
c. 240 physicians stated they do not have this type of patient.
d. 67 physicians stated they do not have this type of patient.
e. 74 physicians stated they do not have this type of patient.

TABLE 1—Multiple Linear Regression Model of Physician STD/HIV Risk Questioning Predictors[a]

Predictor Variable	Beta[c]	Standard Error	*t*	*P*
Intercept	—	—	—	—
Foreign medical school graduate	2.24	0.60	3.73	.0002
Female	1.44	0.60	2.40	.02
Four or more HIV-positive patients	0.79	0.32	2.46	.01
Uncomfortable discussing sex with patients[b]	−1.82	0.32	5.61	.0001
Perceives STD/HIV as a big problem	0.87	0.21	4.19	.0001
HIV prevention confidence[b]	0.32	0.05	6.17	.0001

Note. $R^2 = 0.17$, $F = 26.12$, $df = 6$, $P < .0001$.
[a]Total physician sample = 961; regression physician sample = 849.
[b]Multiple item scaled score.
[c]Unstandardized regression coefficient

below the rates called for by the Public Health Service's Year 2000 objectives.[2] Although the majority of physicians reported regularly asking new patients about their history of STDs during initial general health examinations, less than half of physicians regularly asked about sexual preference, and less than one third of physicians regularly asked about anal or oral sexual practices or number of sexual partners. These risk assessment rates may even be inflated given that physician self-reported practice has been observed to be greater than observed practice.[9-11]

Whether or not a physician reported asking patients about their sexual practices varied widely by patient type. While a majority of study physicians reported regularly asking about the sexual practices of patients in traditional HIV risk groups—homosexual or bisexual men and IV drug users—they infrequently reported regularly assessing the sexual practices of other patients, including single men and women. Therefore, many physicians reported obtaining information about sexual practices only if they already knew a patient was in a high-risk group. If the patient does not bring up the topic of AIDS[12] or his or her high risk STD/HIV status, it appears unlikely that the physician will address the patient's sexual risk of STD/HIV infection.

Three educationally immutable variables were weakly associated with more STD/HIV risk questioning in this study: being a female physician, having cared for four or more HIV-positive patients, and being a foreign-trained physician. These associations are not particularly useful in designing effective CME programs. Three educationally mutable variables were more strongly associated with STD/HIV risk questioning: physician perception that STD/HIV is a big problem in their practice; physician comfort with discussing various

patients', not just homosexual patients',[3] sexual practices; and physician HIV prevention confidence. These potentially mutable factors suggest mechanisms through which physician practice can be improved. Experiential training in sexual questioning[13] may provide physicians with exposure to and skills for assessing patients' personal risk behavior. Such training may decrease feelings of discomfort, increase perceptions about the magnitude of the STD/HIV problem in ones' practice, and increase HIV prevention confidence when discussing patients' sexual practices. At the least, such training may help physicians appropriately handle these barriers. Hence, this study suggests that experiential training for discussing patients' sexual practices may be important in increasing physician STD/HIV risk assessment of all patients. □

Acknowledgments

This study was supported by grants R18 A124403 and R18 A124403-04 from the National Institute for Allergy and Infectious Diseases and a donation from the Gannett Foundation.

Portions of this paper were presented at the 118th Annual Meeting of the American Public Health Association in New York City, October 2, 1990.

References

1. US Preventive Services Task Force. *Guide to Clinical Preventive Services.* Baltimore, Md: Williams & Wilkins; 1989.
2. US Department of Health and Human Services. *Healthy People 2000.* Washington, DC: US Government Printing Office; 1991. US Dept of Health and Human Services publication (PHS)91-50212.
3. Lewis CE, Freeman HE. The sexual history-taking and counseling practices of primary care physicians. *West J Med.* 1987;147:165–167.
4. Lewis CE, Freeman HE, Corey CR. AIDS-related competence of California's primary care physicians. *Am J Public Health.* 1987;77:795–799.
5. Fredman L, Rabin DL, Bowman M, et al. Primary care physicians' assessment and prevention of HIV infection. *Am J Prev Med.* 1989;5:188–195.
6. Gonzalez-Willis A, Rafi I, Boekeloo B, et al. Using simulated patients to train physicians in sexual risk assessment and risk reduction. *Acad Med.* 1990;65(9,suppl): S7–S9.
7. SAS Institute Inc. *SAS/STAT Guide for Personal Computer.* Version 6 Ed. Cary, NC: SAS Institute Inc.; 1987.
8. Lewis CE, Montgomery K. The AIDS-related experiences and practices of primary care physicians in Los Angeles: 1984–89. *Am J Public Health.* 1990;80: 1511–1513.
9. Gemson DH, Colombotos J, Elinson J, Fordyce J, Hynes M, Stoneburner R. Acquired immunodefiency syndrome prevention knowledge, attitudes, and practices of primary care physicians. *Arch Intern Med.* 1991; 151:1102–1108.
10. Norman GR, Neufelf VR, Walsh A, Woodward CA, McConvey GA. Measuring physicians' performance by using simulated patients. *J Med Educ.* 1985;60:925–934.
11. Hoppe RB, Farquhar LJ, Henry R, Stoffelmayr B. Residents' attitudes toward and skills in counseling: using undetected standardized patients. *J Gen Intern Med.* 1990;5:415–420.
12. Gerbert B, Maguire BT, Coates TJ. Are patients talking to their physicians about AIDS? *Am J Public Health.* 1990;80:467–468.
13. Lewis CE. Sexual practices: are physicians addressing the issues? *J Gen Intern Med.* 1990;5(suppl):S78–S81.

Changes in AIDS Case Reporting after Hospital Site Visits

Daniel Fife, MD, James McAnaney, and M. Abid Rahman, PhD

ABSTRACT

In an effort to improve AIDS case reporting, site visits (meetings with hospital staff to encourage reporting) were made to all Philadelphia hospitals. Comparisons of hospitals visited during a 7-week period with hospitals not visited during that period indicated that the site visits were followed by a marked increase in case reports. No similar increase was observed at the comparison hospitals. The increased reporting was accompanied by an increased lag time from diagnosis to report, suggesting that the additional reports at visited hospitals were the result of the identification of previously missed cases rather than a speedup of reporting. Cases reported after the visits were more likely to have white-collar occupations or private medical insurance than were those reported before the visits. (*Am J Public Health.* 1991;81:1648–1650)

Introduction

It is well documented that acquired immunodeficiency syndrome (AIDS) case reporting is incomplete,[1–3] and that many unreported cases can be identified through reviews of death certificates,[4–7] hospital discharge records,[8–10] and outpatient records.[11] In Oregon, a state with relatively few AIDS cases, Modesett recently demonstrated that AIDS case reporting was improved by reminders to hospitals and physicians in addition to reviews of hospital discharges and death certificates.[12]

We sought to determine whether reminders to hospitals also improved AIDS case reporting in a large city with many AIDS cases, and (if such reminders increased AIDS case reporting) to determine whether they did so by bringing to immediate attention cases that might otherwise have been reported in the future or by bringing to attention cases that were previously missed. We differentiated between these two possibilities by examining the lag time from diagnosis to report. In the first situation the lag time should decrease after the reminders (because cases were being reported sooner rather than later). In the second situation the lag time should remain constant or increase (because a new population of cases or a backlog of unreported cases was coming to light).

Methods

In August 1989, Philadelphia's commissioner of health wrote to each of the city's 46 hospitals requesting that the

The authors are with the AIDS Activities Coordinating Office, Philadelphia Department of Health.

Requests for reprints should be sent to Daniel Fife, MD, AIDS Activities Coordinating Office, Philadelphia Department of Health, 1220 Sansom Street, Philadelphia, PA 19107.

This paper was submitted to the journal November 29, 1990, and accepted with revisions May 6, 1991.

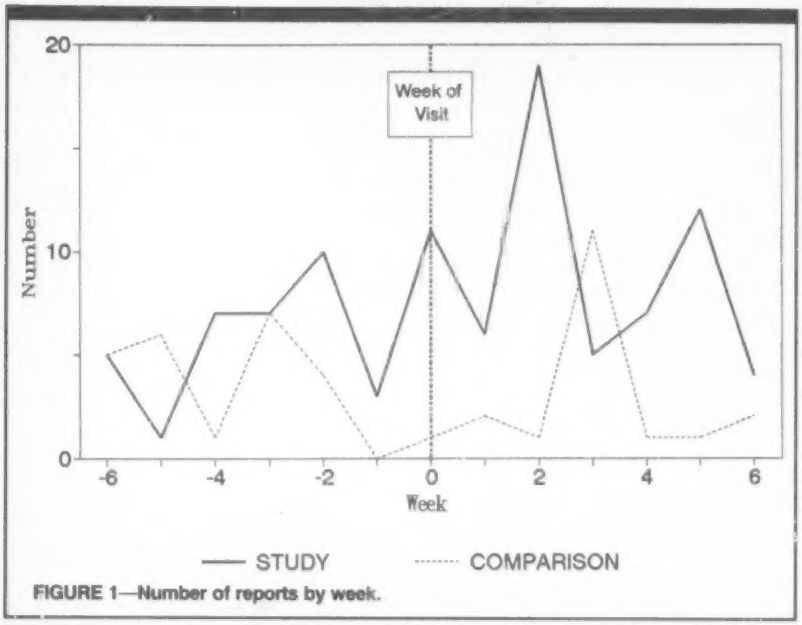


FIGURE 1—Number of reports by week.

TABLE 1—AIDS Case Reports and Mean Lag Time by Hospital Type and Time Period

	6 Weeks Previsit	6 Weeks Postvisit[a]
AIDS case reports, No. (%)		
Study hospitals	33 (58.9)	60 (77.9)[b]
Comparison hospitals	23 (41.1)	17 (22.1)[b]
Total	56 (100.0)	77 (100.0)
Mean lag from diagnosis to report (days)		
Study hospitals	106.4	220.0[c]
Comparison hospitals	162.8	143.3

[a]The week of the visit was included in the postvisit period.
[b]The percentage of AIDS cases identified from the hospitals visited increased significantly ($P < .05$) from the 6 weeks before the visit to the 6 weeks after the visit.
[c]The mean lag time from case diagnosis to case report increased significantly at the hospitals visited when comparing the 6 weeks before the visit with the 6 weeks after the visit ($P = .01$).

administrator meet with Department of Health (DOH) staff to review and improve AIDS case-finding procedures. All hospitals were visited between September 1989 and March 1990, the schedule being determined by the availability of the relevant hospital staff (the administrator and the infection control director). During the 7 weeks beginning January 8, 1990, 21 hospitals were visited. For the present study, these 21 hospitals were considered study hospitals. To provide a comparison hospital for each study hospital, the study hospitals were ordered according to the previous year's number of reported AIDS cases. The remaining Philadelphia hospitals were ordered similarly. A comparison hospital was assigned to each study hospital by pairing the ordered lists.

Reports from each study and comparison hospital pair were grouped by week of receipt relative to the date of the visit to the study hospital. The numbers of reports during the 6 weeks before the visit and the 6 weeks after the visit were compared using a chi-square test, and the mean lag times from diagnosis to report were compared using an unpaired t test (Table 1). Philadelphia collects data on the medical insurance, occupation, and employment status of each reported AIDS case. Medical insurance was grouped as private (Blue Cross, etc.), public (Medicaid, etc.), unknown, or none. Occupation was grouped as white-collar (professional, clerical, etc.) or blue-collar (crafts, labor, etc.).

Visits by DOH staff lasted 30 to 60 minutes and involved a brief discussion of the importance of AIDS case reporting, a detailed review of the hospital's case identification and reporting procedures, and a review of the annual number of reports from that hospital and the mean lag time for those reports. Any problems that were identified were discussed. Most of these discussions led to mutually agreeable changes. The visitors showed maps tracing the increase of AIDS in Philadelphia, distributed pocket cards summarizing the diagnostic criteria for AIDS, thanked the hospital staff for their work so far and for taking the time to meet, and offered to provide feedback on future reporting through lists or summary statistics.

Results

After the visits, the number of case reports increased at the study hospitals and decreased at the comparison hospitals. The proportion of reports coming from the study hospitals increased significantly. The mean lag time for reports from the study hospitals increased significantly. The mean lag time for reports from the comparison hospitals did not increase (Table 1). The increase in the number of reports from the study hospitals began slightly before the site visits (Figure 1).

Comparisons of the cases reported during the 6 weeks before the visit and those reported during the 6 weeks after the visit revealed that, at the study hospitals, there were significant increases in the proportion with white-collar occupations (4/27 [15%] vs 19/48 [41%]; $P < .05$) and private medical insurance (10/33 [30%] vs 26/48 [54%]; $P < .05$). However, at the comparison hospitals there was little change in the proportion with white-collar occupations (4/23 [17%] vs 3/16 [19%]; $P > .05$) or private medical insurance (5/23 [22%] vs 3/16 [19%]; $P > .05$). At both the study and comparison hospitals there were nonsignificant increases in the proportion of cases that reported homosexual contacts and the proportion who were White, and no significant changes in the proportion by gender.

Discussion

The present study provides evidence that hospital site visits are an effective way to increase AIDS case reporting in an urban area with high AIDS rates, and that the additional reports do not simply reflect earlier notification of cases that would have been reported later. The increase in the number of reports from the hospitals

appeared to begin slightly before the site visits. Visits were scheduled several weeks in advance, and it is possible that case finding received additional attention in anticipation of the visit.

The differences in occupation and insurance status between the cases reported before the visits and those reported afterward suggest that it is more difficult to identify AIDS patients with white-collar occupations and private insurance than those in other groups. This possibility deserves further investigation. If AIDS case reporting is biased, with underascertainment of those who are employed in white-collar occupations or have private medical insurance, then current statistics based on reported AIDS cases will overestimate the extent to which AIDS cases are concentrated among socially marginal groups and underestimate the extent to which they occur in other groups.

Relatively little is known about the reasons why cases fail to be reported. In the course of our visits, several administrators commented that we appeared to be marketing case reporting. From that perspective, we could summarize the present study by saying that hospital site visits are an effective way to impress the idea that case reporting deserves space on the medical caregiver's crowded priority list. □

Acknowledgment

This work was partially supported by Centers for Disease Control grant U62/CCU30 6217-01-1.

References

1. United States General Accounting Office. *AIDS Forecasting: Undercount of Cases and Lack of Key Data Weaken Existing Estimates.* Gaithersburg, Md: US General Accounting Office; 1989. US GAO document GAO/PEMD-89-13.
2. Buehler J, Berkelman R, Stehr-Green J, Leary L. Completeness of AIDS surveillance, United States. In: Sixth International Conference on AIDS; June 20–24, 1990; San Francisco, Calif. Abstract Th.C.698.
3. Greenberg AE, Evans R, Lopez E, Thomas P. Completeness of AIDS case reporting and the spectrum of HIV-related disease among outpatients in New York City. In: Sixth International Conference on AIDS; June 20–24, 1990; San Francisco, Calif. Abstract Th.C.708.
4. Chamberland ME, Allen JR, Monroe JM, et al. Acquired immunodeficiency syndrome in New York City: evaluation of an ac[illegible]veillance system. *JAMA.* [illegible]83–387.
5. H[illegible]M, Starcher ET, Morgan WM, et al. Review of death certificates to assess the completeness of AIDS case reporting. *Public Health Rep.* 1987;102:388–391.
6. Conti L, LaLota M, Lieb S, Gill J, Sims J, Witte JJ. Diagnosis at death: an indicator of failure or success? In: Sixth International Conference on AIDS; June 20–24, 1990; San Francisco, Calif. Abstract Th.C.689.
7. McCormick A. Undiagnosed HIV-positive deaths. In: Sixth International Conference on AIDS; June 20–24, 1990; San Francisco, Calif. Abstract Th.C.753.
8. Voght RL, Clark SW, Kappel S. Evaluation of the state surveillance system using hospital discharge diagnoses, 1982–1983. *Am J Epidemiol.* 1986;123:197–198.
9. Conway GA, Colley-Niemeyer B, Pursley C., et al. Underreporting of AIDS cases in South Carolina, 1986 and 1987. *JAMA.* 1989;262:2859–2863.
10. O'Keith G, Froelicher E, Garza A, Lieb L, Eastman C, Williams A. Using International Classification of Diseases—9th Revision (ICD-9) coded hospital discharge data to augment AIDS surveillance. In: Sixth International Conference on AIDS; June 20–24, 1990; San Francisco, Calif. Abstract Th.C.735.
11. Modesitt S, Espenlaub C, Klockner R, Flemming D. AIDS cases diagnoses as outpatients. In: Sixth International Conference on AIDS; June 20–24, 1990; San Francisco, Calif. Abstract Th.C.736.
12. Modesitt SK, Hullman S, Fleming D. Evaluation of active versus passive surveillance in Oregon. *Am J Public Health.* 1990;80:463–464.

Patterns of Breast and Cervical Cancer Screening at Three Public Health Centers in an Inner-City Urban Area

Steven Whitman, PhD, David Ansell, MD, Loretta Lacey, RN, DrPH, Edwin H. Chen, PhD, Nyambi Ebie, MD, Jade Dell, MA, and Clyde W. Phillips, MD

ABSTRACT

In an effort to examine breast and cervical cancer screening patterns among poor African-American urban women, medical records were abstracted at three public health centers located in the inner city of Chicago. The proportions of eligible women at these three centers who received Pap smears, breast examinations, and mammograms were computed. These proportions were notably low and differed significantly among the three centers. Because the literature is now suggesting that an appropriate sequence best defines adequate screening, sequences of screenings were also determined and were found to be lacking. All of these screening histories fall far below the screening objectives set by the National Cancer Institute for the year 2000. This information suggests that interventions are needed that will help health centers serving poor women to deliver more frequent cancer screening.(*Am J Public Health.* 1991;81:1651–1653)

Introduction

African-American women have a higher incidence of cervical cancer and higher mortality rates and lower survival rates than White women.[1] Furthermore, although White women have a higher incidence of breast cancer, the prognosis for African-American women is worse. For example, the ratio of breast cancer mortality to incidence is higher for African-American women (38%) than for White women (32%). Five-year survival rates (63% vs 75%) demonstrate a similar difference.[2] These differences appear to be due in large part to the fact that African-American women with breast cancer are diagnosed at a considerably later stage, even when relevant variables are stratified.[3,4] Such racial differences are particularly troublesome because it now is generally accepted that appropriate screening can reduce mortality from breast and cervical cancer.[5–8]

There are only a few published reports documenting breast and cervical cancer screening patterns in the United States. We have not been able to locate any study that presents specific information on screening patterns for poor urban African-American women. To present such information is our purpose in this paper. We use data from Chicago Department of Health centers to derive the percentages of women who receive Pap smears, breast examinations, and mammograms for selected time intervals and sequences of screens. We then compare our findings with those in the literature and discuss the policy implications that result from this comparison.

Methods

Setting

The data were taken from three public health centers located on the south side of Chicago in very poor neighborhoods consisting almost entirely of African-American people. Virtually all of the patients attending these centers are African-American and about 75% are female.

Data Collection

We abstracted data from medical records to examine the screening histories of African-American women who attended these clinics between February and April 1989. A code manual and code sheet were produced for this purpose and thoroughly pretested. Charts were abstracted for all eligible African-American women who had made an appointment (even if it was not kept) or who walked in and were seen (even if there was no appointment) during the clinic session prior to coding.

The definitions of eligibility, for the purposes of this study, were as follows:

Pap smear	age 40 or over with no hysterectomy
Breast examination	age 40 or over
Mammogram	age 50 or over

Each chart that was reviewed could thus contain information about a woman eligible for no screens, one screen, two screens, or three screens. We were interested not only in when a woman had received her last screening examination, but in what the sequence of screens had been.

The authors are with the Avoidable Mortality from Cancer Project, Epidemiology and Biostatistics Program and Community Health Sciences Program, School of Public Health, University of Illinois at Chicago. Steven Whitman and Jade Dell are also with the Center for Urban Affairs and Policy Research, Northwestern University, Evanston, Ill; David Ansell is also with the Department of Medicine, Cook County Hospital, Chicago, Ill; and Clyde W. Phillips is also with the Cancer Control Program, Chicago Department of Health.

Requests for reprints should be sent to Steven Whitman, PhD, Avoidable Mortality from Cancer Project (M/C 922), School of Public Health, University of Illinois at Chicago, Chicago, IL 60680.

This paper was submitted to the journal October 30, 1990, and accepted with revisions April 24, 1991.

TABLE 1—Percentage of Eligible African-American Women Who Received Screening Examinations at Three Health Centers (A,B,C), by Interval of Screening (ns in parentheses)

Interval, y	Center			Significance[a]		
	A	B	C	A,B	A,C	B,C
			Pap smear			
1	12 (284)	39 (189)	33 (211)	<.001	<.001	NS
2	23 (263)	68 (168)	68 (209)	<.001	<.001	NS
5	59 (189)	81 (115)	83 (159)	<.001	<.001	NS
10	81 (74)	84 (56)		NS	—	—
			Breast examination			
1	6 (454)	36 (352)	12 (369)	<.001	<.005	<.001
2	13 (415)	71 (311)	40 (367)	<.001	<.001	<.001
5	40 (290)	84 (212)	72 (272)	<.001	<.001	<.005
10	76 (121)	87 (101)		<.05	—	—
			Mammogram			
1	3 (394)	17 (303)	24 (337)	<.001	<.001	<.05
2	6 (367)	32 (269)	46 (336)	<.001	<.001	<.001
5	12 (262)	54 (189)	60 (257)	<.001	<.001	NS
10	12 (113)	48 (91)		<.001	—	—

Note. Eligible women were defined as follows: for Pap smear, age 40 or over, with no hysterectomy; for breast examination, age 40 or over; for mammogram, age 50 or over.
[a]NS = not significant.

TABLE 2—Percentage of Eligible African-American Women Who Received Screening Examinations at Three Health Centers (A,B,C), by Number Received within Previous 5 Years

Number of Exams	Center			Significance[a]		
	A	B	C	A,B	A,C	B,C
	Pap Smear					
	(n = 192)	(n = 141)	(n = 163)			
<2	72	29	47			
2	12	20	22			
>2	16	51	31	<.001	<.001	<.005
	Breast Examination					
	(n = 291)	(n = 211)	(n = 272)			
<2	87	37	64			
2	7	26	24			
>2	6	37	12	<.001	<.001	<.001
	Mammogram					
	(n = 262)	(n = 188)	(n = 257)			
<2	98	84	77			
2	2	12	16			
>2	<1	4	7	<.001	<.001	NS

Note. Eligible women were defined as follows: for Pap smear, age 40 or over, with no hysterectomy; for breast examination, age 40 or over; for mammogram, age 50 or over.
[a]Significance results were computed only for percentages for >2 screens. NS = not significant.

We therefore abstracted data for the previous 10 years at Centers A and B, and the previous 5 years at Center C depending on how long the woman had been attending the clinic. Coding was conducted for 6 days at each of the centers. Sample sizes for Pap smears are smaller than for breast examinations or mammograms because so many women (over 40%) at these centers had had a hysterectomy.

Data Analysis

All analyses are based on comparisons of proportions of women screened. A two-tailed *t* test of proportions was used, along with a significance level of 5%.

Results

Table 1 presents the percentage of women screened at each of the three centers and the corresponding sample sizes. Only 12% of eligible women at Center A had received a Pap smear within the previous year, compared with 39% at Center B and 33% at Center C. The percentages for Centers B and C do not significantly differ ($P > .05$), but both centers performed a significantly larger percentage of Pap smears ($P < .001$) than Center A. Percentages of women who had had a Pap smear within the previous 2 years, 5 years, and 10 years are also shown in Table 1, along with the corresponding significance results. About 18% of the women at Centers A and B had not had a Pap smear within the past 10 years. Table 1 also presents results for breast examinations and mammograms. The patterns for these two screening procedures are similar to those for Pap smears.

We also analyzed the sequence of screening. Table 2 shows the percentage of women who had been attending these clinics for at least 5 years and who had had fewer than two, two, and more than two Pap smears, breast examinations, and mammograms during the previous 5 years. Most recommendations for these procedures for women in these age groups suggest annual screening.[9-11] Thus, let us conservatively define "appropriate screening" as more than two screens in the past 5 years. Using such a definition, only 16% of the women at Center A were being appropriately screened for cervical cancer, compared with 51% at Center B

and 31% at Center C. As indicated in Table 2, all of these percentages differ significantly from one another. Similar patterns exist for breast examinations and mammograms, also presented in Table 2.

TABLE 3—Average Number of Visits per Year by Eligible African-American Women to Three Health Centers

	Center A	Center B	Center C
Average number of visits	35.6	35.7	26.9
Average number of years attending	7.1	7.2	7.5
Average number of visits per year	5.0	5.0	3.6

Note. Eligible women were defined as follows: for Pap smear, age 40 or over, with no hysterectomy; for breast examination, age 40 or over; for mammogram, age 50 or over.

Discussion

Most of the observed differences in the percentages of eligible women receiving a screen in a given interval or a sequence of screens in a given interval are statistically significant. Although we performed many significance tests, thus somewhat altering the true significance levels of the tests, the trends are unmistakable. Center A was screening women at a far lower rate than Centers B and C; to a lesser extent, Center C was doing significantly worse than Center B. We do not yet have an adequate explanation for such differences.

In order to try to understand the causes of these differences, we examined the demographic characteristics of the clinics. In another study, we found that the women served by Centers A and B were remarkably well matched on virtually all demographic characteristics, as well as on factors related to knowledge and attitudes.[12] Center C was not part of the latter investigation, but we know from unpublished data that the women served by this Center are demographically well matched with the women served by Centers A and B.

There is now some excellent literature, almost all of it from Europe, that suggests that the hallmark of good screening is a sequence of multiple negative tests within a "short interval."[6,8,13] Much of this literature is methodologically sound and presents convincing evidence that effective screening depends on such a sequence. Unfortunately, these studies present varying results with respect to what this optimal number may be and how long this "short interval" is. We can see from Table 2 that a very large percentage of the women at these centers are not receiving "appropriate" cancer screening for cervical cancer, when this definition is based on a standard of more than two screens in 5 years. The results are even more sobering for breast examinations and mammograms.

These results take on additional importance in light of the average number of clinic visits for the women in our sample (Table 3). The women attending Centers A and B averaged 5 visits per year, while the women attending Center C averaged 3.6 visits per year. These women have frequent access to health care resources, yet still do not receive adequate cancer screening.

The National Cancer Institute has set as a goal for the year 2000 that 80% to 90% of women will have Pap smears every 3 years and that 80% of women between the ages of 50 and 70 will receive annual breast examinations and mammograms.[14] The data presented in Tables 1 and 2 suggest that these goals are not even being approximated for the women attending the health centers in our study.

To our knowledge, this is the first study that presents an image of breast and cervical cancer screening derived from medical records for poor African-American women who most need cancer screening.[15,16] It also appears to be the first study from the United States that presents sequences of such screenings. Based on these findings, we conclude that the women in our sample, all of whom were attending public sector clinics, are not receiving appropriate cancer screening. This is a situation that demands remediation. □

References

1. Baquet C, Ringen K. Cancer control in Blacks: Epidemiology and NCI program plans. In: *Advances in Cancer Control: Health Care Financing and Research.* eds: LE Mortenson, PF Engstrom, PN Anderson, New York: Alan R. Liss, Inc.;1986: 215–227.
2. National Cancer Institute. *Seer Program: Cancer Incidence and Mortality in the United States, 1973–1981* (revised). Washington, DC: US Govt Printing Office; November 1984. DHHS publication NCI 85-1837.
3. Farley TA, Flannery JT. Late-stage diagnosis of breast cancer in women of lower socioeconomic status: public health implications. *Am J Public Health.* 1989;79: 1508–1512.
4. Satariano WA, Belle SH, Swanson GM. The severity of breast cancer at diagnosis: a comparison of age and extent of disease in Black and White women. *Am J Public Health.* 1986;76:779–782.
5. Shapiro S. The status of breast screening: a quarter of a century of research. *World J Surg.* 1989;13:9–18.
6. Van der Graaf Y, Klinkhamer PJJM, Vooijs GP. Effect of population screening for cancer of the uterine cervix in Nijmegen, The Netherlands. *Prev Med.* 1986;15:582–590.
7. Tabar L, Gad A, Holmberg LH, et al. Reduction in mortality from breast cancer after mass screening with mammography. *Lancet.* April 13, 1985:829–832.
8. Creasman WT, Fowler WC, Homesley HD. Screening for cervical cancer: 1982—an update. *N Carolina Med J.* November 1982:771–773.
9. Smart CR. Critique of the early cancer detection guidelines of the U.S. preventive services task force and the National Cancer Institute. *Mayo Clinic Proc.* 1990;65:892–898.
10. Think DJ. Change in American Cancer Society checkup guidelines for detection of cervical cancer.CA—*J Clin.* 1988;38:127–128.
11. US Prevention Service Task Force. *Guide to Clinical Prevention Services. An Assessment of the Effectiveness of 169 Interventions.* Baltimore, Md: Williams & Wilkins Co; 1989.
12. Lacey L, Ansell D, Whitman S, Chen E, Ebie N, Phillips CW. Knowledge and attitudes about breast and cervical cancer among Black women. Presented at the American Public Health Association Annual Meeting; October 25, 1989; Chicago, Ill.
13. Fink R, Shapiro S, Roester R. Impact of efforts to increase participation in repetitive screenings for early breast cancer detection. *Am J Public Health.* 1972;62:328–336.
14. National Cancer Institute. *Cancer Control Objectives for the Nation: 1985–2000.* Washington, DC: US Govt Printing Office; 1986. DHHS publication NCI 86-2880.
15. Hayward RA, Shapiro MF, Freeman HE, Corey CR. Who gets screened for cervical and breast cancer? *Arch Intern Med.* 1988;148:1177–1181.
16. The NCI Breast Cancer Screening Consortium. Screening mammography: a missed clinical opportunity? *JAMA.* 1990;264:54–58.

The Impact of Vitamin A Supplementation on Morbidity: A Randomized Community Intervention Trial

Mutee H. Abdeljaber, MD, DrPH, Arnold S. Monto, MD, Robert L. Tilden, PhD, M. Anthony Schork, PhD, and Ignatius Tarwotjo, PhD

ABSTRACT

A randomized community trial investigated the impact of vitamin A supplementation on symptoms of respiratory and enteric infections among Indonesian children aged 1 to 5 years. A sample of 450 villages had been randomized into 229 villages that received vitamin A distribution and 221 that were observed as controls. The prevalence rates of cough, fever, and diarrhea were reduced in the treatment villages compared with the control villages by 8%, 5%, and 11%, respectively, but the reduction was not statistically significant. These results indicate that vitamin A supplements did not produce a substantial reduction in these symptoms, in spite of a reported reduction in all-cause mortality. (*Am J Public Health.* 1991;81:1654–1656)

Introduction

Many have hypothesized that vitamin A deficiency plays a major role in mortality and morbidity from respiratory and enteric infections.[1–5] In an Indonesian study, preschool children with vitamin A deficiency (measured by xerophthalmia) were at a twofold higher risk of developing respiratory tract infections and a threefold higher risk of developing diarrhea than their peers without xerophthalmia.[6] An Indian study reported that children with mild xerophthalmia were again at a twofold risk of developing respiratory tract infections, but were not at greater risk of developing diarrhea than controls.[7] In a later Indonesian study, vitamin A supplements twice a year reduced mortality by 34% in the treated group compared with controls, but the impact of the supplements on morbidity has never been analyzed.[8] The current study performs such an analysis, examining the effect of vitamin A supplementation in reducing the frequency of symptoms of respiratory and enteric infections.

Methods

The data had been collected by the Aceh study group, in Aceh province at the northern tip of Sumatra, Indonesia, to investigate the impact of vitamin A supplements on xerophthalmia, mortality, and growth.[8,9] The Aceh study consisted of a sample of 450 villages, randomized into 229 villages (15 101 resident children aged 1 to 5 years), that received vitamin A distribution and 221 villages (13 760 resident children) that were observed as controls. Each village in the two groups was visited twice, first for the baseline survey and second for the follow-up survey 1 year later. Capsules of vitamin A, containing 200 000 international units were distributed by a trained volunteer to preschool children in the treatment villages at the time of the first visit and 6 months later.

The first survey was conducted between September 1982 and August 1983, by a trained team consisting of an ophthalmologist and five enumerators. Households having children under 60 months of age were identified and asked about history of their having cough and fever, and diarrhea, in the previous week. Cough and fever that occurred and lasted for at least 24 hours in the previous week were used to indicate respiratory infection. Diarrhea, defined as having loose or watery stools for four or more times per day in the previous week, was used to indicate enteric infection. At baseline, socioeconomic data were collected concerning paternal age, education, and occupation, as well as house sanitation and source of drinking water. Children were examined for xerophthalmia according to standard criteria.[10] Follow-up visits were made 9 to 13 months after the baseline. The same team visited each village, reexamined the children's eyes, and recorded the number of deaths and the occurrence of diarrhea, cough, and fever during the previous week. Since randomization was carried out on the village level, data were aggregated by village (means and percentages of the appropriate population), instead of by child.

Mutee H. Abdeljaber and Arnold S. Monto are with the School of Public Health, Arnold S. Monto and Robert L. Tilden are with the Department of Population Planning and International Health, M. Anthony Schork is with the Department of Biostatistics, all at the University of Michigan, Ann Arbor. Ignatius Tarwotjo is with the Ministry of Health, Republic of Indonesia Directorate General of Community Health, Jakarta, Indonesia. All authors are members of the Aceh, Indonesia Study Group.

Requests for reprints should be sent to Arnold S. Monto, MD, School of Public Health, University of Michigan, Ann Arbor, MI 48109.

This paper was submitted to the journal July 12, 1990, and accepted with revisions April 24, 1991.

Results

Table 1 shows that the two groups of villages were comparable in baseline prevalence of cough, fever, diarrhea, and xerophthalmia ($P > .05$). The two groups were also comparable for other baseline demographic and socioeconomic variables that might be related to the transmission of infections. The age and sex distribution of children and the village size in the two groups were similar.[8,9] Table 2 shows the prevalence rates of cough, fever, and diarrhea in the two study groups before and after vitamin A intervention. Data for the two surveys were collected at the same time of the year by the same team, whose members received standardized training. There was an unexpected reduction in the postintervention prevalence rates in the two study groups compared with the preintervention rates. Postintervention prevalence of cough and fever in the treatment group was 17.0 and 25.0 cases per 100 children, compared with 18.6 and 26.7 cases per 100 children in the control group. Thus, the risk difference was 1.6 per 100 children for cough (95% confidence interval [CI] = −5.7 to 8.9) and 1.7 per 100 children for fever (95% CI = −6.6 to 9.9), and the percentage reduction in the point prevalence in the intervention group was 8.6% for cough and 6.4% for fever. These differences were very small and statistically insignificant ($P > .05$). The prevalence of diarrhea in the intervention group was 4.0 cases per 100 children at the follow-up survey compared with 4.5 cases per 100 children in the control group. The risk difference between the two groups was 0.5 case per 100 children (95% CI = −4.2 to 5.2). The percentage reduction in diarrheal rates in the treatment group was 11.1%, as compared with the control group.

TABLE 1—Comparison of Baseline Characteristics between the Two Groups of Villages: Controls and Treatment

Characteristics	Control (209 villages)	Treatment (219 villages)	*P* Value[a]
Baseline prevalence of symptoms (cases/100 children)			
Cough	31.7	31.7	.97
Fever	46.1	46.2	.96
Diarrhea	7.5	7.4	.90
Xerophthalmia (cases/100 children)	2.30	3.40	.09
Mother's age (mean %)			
<20	3.00	3.50	.75
21–30	48.00	48.00	.68
31–40	35.00	36.00	.48
41–50	11.00	11.00	.48
50+	3.00	1.50	.15
Education			
6 years	82.00	83.00	.42
Source of drinking water (mean %)			
Well	58.10	55.40	.34
Spring	36.40	37.70	.62
River	4.70	6.50	.35
House sanitation (mean %)			
Public pit	74.00	77.00	.11
Private pit	24.00	22.00	.42
Toilet	0.60	0.50	.36
Mean no. of children	63.3	66.8	.40
Distribution of village size (percentile distribution)			
0 to 25	37	40	—
26 to 50	18	17	—
51 to 75	24	25	—
76 to 100	130	136	—

[a]By *t* test.

TABLE 2—Prevalence Rates of Diarrhea, Cough, and Fever before and after Vitamin A Treatment (No. of cases per 100 children) in the Aceh Clinical Trial Study

	Control Group (n = 209)		Treatment Group (n = 219)	
Symptoms of Infection	Pre-treatment	Post-treatment	Pre-treatment	Post-treatment
Cough	31.7	18.6	31.7	17.0
Fever	46.1	26.7	46.2	25.0
Diarrhea	7.5	4.5	7.4	4.0

Discussion

Mortality analysis, using the Indonesian Aceh study, showed that vitamin A supplements to preschool children with marginal vitamin A deficiency reduced mortality by 34%.[8] The present analysis examined the impact of vitamin A supplements on morbidity in terms of cough, fever, and diarrhea. Morbidity should be a more direct outcome of marginal vitamin A deficiency than mortality, which is the outcome of complex health-related variables.[11,12] However, the results demonstrated a statistically insignificant reduction in the rates of cough, fever, and diarrhea.

While the vitamin A intervention did not produce significant differences in morbidity, socioenvironmental factors such as sanitation, water supply, and season were significantly related to these outcomes.[13] This indicates that even though data were aggregated by village, which could potentially attenuate associations, these differences could still be detected.[14] A variety of explanations can be offered for finding, in the same data, a major vitamin A effect on mortality but not on symptoms of the two leading causes of childhood death—respiratory and enteric infections. It is possible that the effect of vitamin A intervention may be on case fatality, with no decrease in frequency and duration of symptoms of respiratory and enteric diseases. In the Philippines, the findings of Tupasi et al., who reported that low socioeconomic status was associated with morbidity and overall malnutrition was associated with mortality, might support this hypothesis.[15] In any event, the current analysis indicates that large-dose vitamin A supplementation is unlikely to produce a rapid and dramatic effect on respiratory and enteric morbidity in the absence of changes in overall malnutrition and socioenvironmental factors. Research currently in progress should clarify the relationship between vitamin A deficiency and mortality and determine whether

these results can be replicated in settings with different socioenvironmental risk factors. □

Acknowledgments

This research was supported by the United States Agency for International Development—Jakarta Mission, PT Richardson-Vicks, the Ford Foundation, UNICEF, the Mobil Oil Corporation, and the Asian Foundation for Prevention of Blindness.

The contributions of the following, additional members of the Aceh Study are acknowledged: A. A. Loedin, Imam Satibi, Sukarno Nur, Henny Neudrawati, Lisa Mele, Daniel Kraushaar, and Hugh Taylor.

We would like to acknowledge Suzzine Pak for her help preparing the manuscript and Margo Goode for her editing skills.

References

1. Handayani T, Mujiani HV, Rhode J. Child mortality in a rural Japanese village: a prospective study. *Int J Epidemiol.* 1983;12:88–92.
2. World Health Organization. Diarrheal Diseases Control Program: an overview of the problem in America. *Weekly Epidemiol Rec.* 1982;57:353–360.
3. Chandra RK. Reduced secretory antibody response to live attenuated measles and poliovirus vaccines in malnourished children. *Br Med J.* 1975;ii:583–585.
4. Cohen BE, Elin RJ. Vitamin A induced non-specific resistance to infection. *J Infect Dis.* 1974;129(5):597–600.
5. Watson RR. *Nutrition, Disease Resistance and Immune Function.* New York, NY: Marcel Dekker; 1984:299–312.
6. Sommer A, Katz J, Tarwotjo I. Increased risk of respiratory disease and diarrhea in children with preexisting mild vitamin A deficiency. *Am J Nutr.* 1984;40:1090–1095.
7. Milton R, Reddy V, Naidu A. Mild vitamin A deficiency and childhood morbidity: an Indian experience. *Am J Clin Nutr.* 1987;46:827–829.
8. Sommer A, Tarwotjo I, Djunaedi E, West K, Loeden R, Mele L. Impact of vitamin A supplementation on childhood mortality. *Lancet.* 1986;1:1169–1173.
9. West K, Djunaedi E, Pandji A, Kusdiono, Tarwatjo I, Sommer A. Vitamin A supplementation and growth: a randomized community trial. *Am J Clin Nutr.* 1988;48:1257–1264.
10. Sommer A. *Nutritional Blindness: Xerophthalmia and Keratomalacia.* Oxford, England: Oxford University Press; 1982.
11. Feachem RC. Vitamin A deficiency and diarrhea: a review of interrelationships and their implications for the control of xerophthalmia and diarrhea. *Trop Dis Bull.* 1987;84(3):R1–R16.
12. Feachem RG. Infections related to water and excreta: the health dimension of the decade. In: Dangerfield BJ, ed. *Water Supply and Sanitation in Developing Countries* (Water Practice Manuals). No. 3. London: Institution of Water Engineers and Scientists; 1983:25–46.
13. Abdeljaber MH. Doctoral thesis, University of Michigan. Impact of Vitamin A Supplements on Childhood Morbidity: A Randomized Community Trial; 1989
14. Greenland S. Limitation of the logistic analysis of epidemiologic data. *Am J Epidemiol.* 1980;110:693–698.
15. Tupasi TE, Velmonte MA, Sanvectores ME, et al. Determinants of mortality and morbidity due to acute respiratory infections. *J Infect Dis.* 1988;157(4):615–623.

Occupational Injuries and Illnesses among Washington State Agricultural Workers

Paul Demers, MS, and Linda Rosenstock, MD, MPH

ABSTRACT

In Washington state 29 451 workers' compensation claims were filed by farm workers between 1982 and 1986. Five percent of claimants were under 18 years of age, including 1.3% under the age of 16. Agricultural workers were found to be at higher risk than other workers in the state for fatal injury (relative risk [RR] = 2.5, 95% confidence interval [CI] = 1.7–3.7), sprains and strains (RR = 1.4, CI = 1.3–1.4), fractures (RR = 2.3, CI = 2.2–2.4), dislocations (RR = 1.9, CI = 1.7–2.2), concussions (RR = 1.9, CI = 1.6–2.3), amputations (RR = 2.5, CI = 2.0–3.1), dermatitis (RR = 4.3, CI = 4.0–4.7), systemic poisoning (RR = 3.4, CI = 3.1–3.7), respiratory disease (RR = 1.7, CI = 1.0–2.9), and tendonitis (RR = 1.2, CI = 1.1–1.4). (*Am J Public Health.* 1991;81:1656–1658)

Introduction

Agricultural employment is among the nation's most hazardous occupations.[1,2] Farm workers have been noted to be at increased risk for musculoskeletal disorders, noise-induced hearing loss, dermatitis, respiratory diseases, and cancer.[1,3–5] The most widely recognized hazards of farm work are pesticides and agricultural machinery, but farm workers are also exposed to severe climatic conditions, physical fatigue and stress, parasites, infectious diseases, toxic plants, and chemicals needed to maintain farm machinery.[5,6]

We reviewed workers' compensation claim data for farm workers in the state of Washington between 1982 and 1986 in order to characterize the nature of work-related illnesses and injuries among farm workers and compare their rates to those of nonagricultural workers.

Methods

The state supplied the investigators with computer tapes containing data on all claims filed in Washington between January 1, 1982, and December 31, 1986. The nature and source of injury and illness claims were derived from the first report received at the time the claim was filed, and claim status was determined as of April 1988. Using hours submitted to the state by employers for billing purposes, the number of full-time equivalent (FTE) agricultural workers was estimated. In or-

The authors are with the Occupational Medicine Program, University of Washington, Seattle.

Requests for reprints should be sent to Linda Rosenstock, MD, MPH, Occupational Medicine Program, University of Washington, Harborview Medical Center, ZA-66, 325 9th Avenue, Seattle, WA 98104.

This paper was submitted to the journal September 24, 1990, and accepted with revisions April 2, 1991.

TABLE 1—Workers' Compensation Claim Status: Washington State Farm Workers, 1982 to 1986

Claim Status	Total
Medical bills only	19 994
Time loss paid	7375
Total permanent disability	27
Fatality	18
Other*	305
Rejected	1732
Total	29 451

*Kept on salary, loss of earning power, or no decision as of the time the data were generated.

der to calculate rates for nonagricultural workers, data from the state employment security commission were used. Relative risks were calculated by comparing the rate for agricultural workers to the rate for all other workers in the state. Confidence intervals were calculated using Taylor series methods.[7]

Results

A total of 29 451 claims were filed by farm workers between 1982 and 1986. The status of all claims as of 1988 is given in Table 1. Table 2 shows the number of accepted claims by the type of injury or illness. The claims rate for agricultural workers was almost 50% higher than for nonagricultural workers, with substantial excesses for each of the categories examined except hearing loss (Table 3).

Between 1982 and 1986 only 53 claims were filed for organophosphate pesticide (OP) related illnesses, with 11 (21%) rejected and only 3 (6%) resulting in compensable time loss. Of the 54 claims filed for pesticides other than organophosphates, 9 (17%) were rejected and only 3 (6%) resulted in compensable time loss. During the study period there were 148 dermatitis claims, and 106 systemic poisoning claims were attributed to "other and unspecified chemicals."

Although 74% of all claimants were between the ages of 18 and 40, 5% were under 18 years of age and 1.3% were under 16. Ninety-four percent of these claims were accepted and 20% had compensable time loss. One claim was accepted for a childhood fatality: a 15-year-old orchard worker died from multiple injuries following a work-related motorcycle accident. Other serious injuries for which claims were accepted for children included 7 amputations, 90 fractures, 11 dislocations, and 8 concussions. These serious injuries represented 8.8% of all claims accepted for children vs 7.1% of all claims filed by adults.

Discussion

We found agricultural workers in the state of Washington to be at higher risk than other workers for both occupational injuries and illnesses. Approximately 5% of the claims examined by this study were filed for children. It is not possible to determine if the higher proportion of serious injuries seen in children compared to adult farm workers was because children are at higher risk or because claims were not filed for less serious injuries. Because of the serious nature of many of these claims, this population should receive special attention by regulatory and health agencies.[8]

We were able to identify relatively few pesticide related claims. This may be due to a lower than anticipated risk or to inaccurate or underreporting. In analyses of California pesticide-related illnesses, Kahn[9] and Blanc et al.[10] found that only a small fraction of cases were reported despite the fact that physicians are required to report pesticide poisonings in that state. A survey of farm workers in Washington State found that only 8% to 15% of those who felt ill after exposure to pesticides sought medical attention.[11] In addition, the relatively large number of claims attributable to other and unspecified chemicals points to the need for better reporting by physicians and more detailed coding by the responsible state agency. The recent enactment of legislation making pesticide poisonings reportable in Washington may provide a better means of surveillance in the future.

This study suffers from an inherent limitation: compensation records are known to be insensitive to the total extent of occupational disease.[9,10,12–14] Despite this limitation, recent investigations have found that internal analyses of compensation data appear to be valid.[15,16] Moreover, in the case of farm workers few other sources of data are available. An additional limitation was the lack of reliable estimates of the number of farm workers in the state during the study period. One advantage of using workers' compensation data is that they are based on hours and allow the calculation of FTE workers. Because of the seasonal nature of farm labor, the use of FTE workers rather than estimates made at one particular time during the year may better represent person-years at risk.

TABLE 2—Accepted Injury or Illness Claims: Washington State Farm Workers, 1982 to 1986

Injury or Illness	Accepted Claims	(%)
All injuries	25 097	(90.6)
Strains and sprains	8991	(32.5)
Lacerations	5353	(19.3)
Contusions	3840	(13.9)
Abrasions	2602	(9.4)
Fractures	1679	(6.1)
Chemical burns	245	(0.9)
Hernias	198	(0.7)
Heat burns	211	(0.8)
Dislocations	187	(0.7)
Concussions	100	(0.4)
Amputations	78	(0.3)
Electric shocks	12	(<0.1)
Multiple injuries	1601	(5.8)
All illnesses	1754	(6.3)
Contact dermatitis	396	(1.4)
Allergic dermatitis	66	(0.2)
Other dermatitis	69	(0.2)
Systemic effects due to toxic materials	403	(1.5)
Bronchitis, asthma and related	38	(0.1)
Upper respiratory	16	(<0.1)
Other systemic poisoning	57	(0.2)
Conjunctivitis	224	(0.8)
Other diseases of the eye	33	(0.1)
Tendonitis	221	(0.8)
Welders flash	93	(0.3)
Peripheral neuropathy	66	(0.2)
Infectious diseases	18	(0.1)
Respiratory disease	12	(<0.1)
Heat strokes	11	(<0.1)
Hearing loss	7	(<0.1)
Freezing	9	(<0.1)
Other illness	13	(<0.1)
Other, ill defined, or unknown	868	(3.1)
Total	27 719	(100)

The increased risk for agricultural workers compared to other workers could be explained by different reporting rather than a real difference between the two populations. This could occur if farm workers were more likely to submit claims or if their numbers are underestimated. Because some farm work is paid for on a piece rate rather than on an hourly basis, farm owners may sometimes be forced to calculate hours based on a formula using

TABLE 3—Comparison of Workers' Compensation Claims Rates between Agricultural and Nonagricultural Workers: Washington State, 1982 to 1986[a]

Type of Claim	Agricultural Workers Rate	Nonagricultural Workers Rate	Rate Ratio (95% Confidence Limits)
All accepted claims	207.84	139.76	1.49 (1.48–1.50)
Sprains and Strains	69.08	49.91	1.38 (1.35–1.41)
Fractures	12.41	5.35	2.32 (2.21–2.43)
Dislocations	1.39	0.73	1.91 (1.66–2.20)
Concussions	0.76	0.40	1.92 (1.59–2.32)
Amputations	0.59	0.23	2.53 (2.04–3.14)
Dermatitis	3.83	0.89	4.32 (3.96–4.71)
Systemic poisoning	3.89	1.15	3.39 (3.11–3.69)
Tendonitis	1.65	1.32	1.25 (1.10–1.42)
Hearing loss	0.05	0.29	0.16 (0.08–0.34)
Respiratory disease	0.09	0.05	1.69 (0.97–2.94)
Fatalities	0.19	0.08	2.50 (1.71–3.66)

[a]Rates per 1000 person years at risk.

the average number of bushels picked per hour or some other productivity estimate. While it is possible that this practice and the general difficulties of enforcing workers' compensation laws in agriculture could lead to some underreporting of hours, a bias in this direction would have to be extremely large to explain the magnitude of the observed differences for most of the outcomes, and substantial fines may be levied against an employer if a claim is filed and no workers' compensation premiums have been paid.

The analysis of workers' compensation claims submitted by farm workers in Washington State has provided a means to examine some aspects of work-related injuries and illnesses in this understudied population. The needs of farm workers should be prioritized for regulatory action and other forms of intervention. □

Acknowledgments

This study was funded through a grant from the Milbank Memorial Fund. We would like to acknowledge the University of Washington Pesticide Health Effects Study Group, including Scott Barnhart, Lucio Costa, William E. Daniell, David Eaton, Sharon Morris, and Sheldon Murphy (deceased). We would also like to thank Pam Hill, Bridget Flory, and Debbie Schlenker from the Washington (state) Department of Labor and Industries, Rollin Reider and Bob Jamieson from the Academic Computer Center of the University of Washington, and Goldie Kleinman for her advice and assistance throughout the study.

References

1. Coye MJ. The health effects of agricultural production: I. the health of agricultural workers. *J Public Health Policy.* 1985;6:349–370.
2. Purschwitz MA, Field WE. Scope and magnitude of injuries in the agricultural workplace. *Am J Ind Med.* 1990;18:179–192.
3. Moses M. Pesticide-related health problems and farm workers. *Am Assoc Occup Health Nurs J.* 1989;37:115–130.
4. Strigini P. On the political economy of risk: farm workers, pesticides, and dollars. *Int J Health Serv.* 1983;12:263–292.
5. May JJ. Issues in agricultural safety. *Am J Ind Med.* 1990;18:121–131.
6. Myers JR. National surveillance of occupational fatalities in agriculture. *Am J Ind Med.* 1990;18:163–168.
7. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiological research: principles and quantitative methods. New York, NY: Van Nostrand Reinhold Co; 1982:351–355.
8. Pollack SH, Landrigan PJ. Child labor in 1990: prevalence and health hazards. *Annu Rev Public Health.* 1990;11:359–375.
9. Kahn E. Pesticide related illness in California farm workers. *J Occup Med.* 1976;18:693–696.
10. Blanc PD, Rempel D, Maizlish N, Hiatt P, Olson KR. Occupational illness: case detection by poison control surveillance. *Ann Intern Med.* 1989;111:238–244.
11. Mentzer M, Villalba B. Pesticide exposure and health: a study of Washington farm workers. Granger, Wash: Evergreen Legal Services, Farm Worker Division; 1988.
12. Discher DP, Kleinman GP, Foster FJ. *Pilot Study for Development of an Occupational Disease Surveillance Method.* Washington, DC: 1975. US Dept of Health, Education, and Welfare publication (NIOSH)75-162.
13. National Research Council. *Counting Injuries and Illnesses in the Workplace—Proposals for a Better System.* Washington, DC: National Academy Press; 1987.
14. Institute of Medicine. *Addressing the Physician Shortage in Occupational and Environmental Medicine.* Washington, DC: National Academy Press; 1991.
15. O'Malley M, Thun M, Morrison J, Toby Mathias CG, Halperin WE. Surveillance of occupational skin diseases using the supplemental data system. *Am J Ind Med.* 1988;13:291–299.
16. O'Malley M, Toby Mathias CG. Distribution of lost-work-time claims for skin disease in California agriculture. *Am J Ind Med.* 1988;13:715–720.

The Effects of Removing a Statutory Barrier to Workers' Compensation for Farm Workers

Paul Demers, MS, and Linda Rosenstock, MD, MPH

ABSTRACT

In 1983 the Washington State Supreme Court ruled unconstitutional a unique exemption that excluded farm workers from workers' compensation if they had not earned $150 in continuous work with the same employer. In order to examine the effect of removing this restriction, workers' compensation data were analyzed to examine the number of farm workers covered, the number of claims submitted, and claim acceptance rates before and after the ruling. Although the 1983 ruling appears to have had a significant impact on the treatment of claims, no increase in utilization of the workers' compensation system was observed. Further intervention is required if the objective of treating farm workers equitably is to be realized. (*Am J Public Health.* 1991;81:1659–1660)

Introduction

Differential treatment by government regulatory agencies has contributed to the plight of farm workers.[1–3] Even in cases where agencies have been mandated to treat farm workers equitably, enforcement is hampered by the lack of staff with training in farm hazards and fluency in Spanish, the seasonal nature of agriculture, and the difficulty in reaching widely scattered and often remote work places. Farm workers have also received differential treatment in regard to workers compensation, and many states still do not cover farm workers.[4] In Washington State prior to 1984, agricultural workers were exempted from coverage if they had not earned $150 in continuous employment with the same employer. Because of this exclusion, farm worker advocates estimate that at any time 20% to 40% of farm workers were ineligible for compensation because of the pattern of frequently changing and intermittent employment and the employers' awareness of this liability caveat. In September of 1983, 7 years after it was legally challenged, the Washington State Supreme Court ruled that the exemption was unconstitutional because it denied equal protection to farm workers by infringing on their right to travel. In October of that same year the state complied with this ruling and began universal coverage for farm workers.

The objective of this study was to explore the effect of the 1983 Washington State Supreme Court ruling expanding farm worker coverage under the state's compensation system.

Methods

The Washington (State) Department of Labor and Industries supplied the investigators with computer tapes containing data on all claims filed in Washington between January 1, 1982, and December 31, 1986. Workers' compensation coverage, utilization of the system, and treatment of farm worker claims were examined before and after the Washington State Supreme Court ruling. Using hours submitted to the state by agricultural employers for billing purposes, the number of full-time equivalent (FTE) agricultural workers was estimated in order to evaluate trends in the number of workers theoretically covered. Two occupational groups were used for comparison: farmers and farm managers (1980 US Census codes 473–477) and unskilled or semi-skilled blue collar workers (703–889).

Results

The number of FTE agricultural workers covered by the system based on premiums paid remained relatively stable during the 5-year period (Table 1). Approximately 2000 more agricultural workers were covered in 1984 than in 1983, a 7% increase. The number of both submitted and accepted farm worker claims peaked in 1983 and declined in each succeeding year. There were fewer claims filed in the last quarter of 1983 (1487) than in the same quarter of either the preceding (1604) or following (1644) years, indicating that the 1983 figures were not distorted by a flood of claims submitted following the Washington State Supreme Court ruling.

Table 2 presents the rejection rates for claims made by farm workers, farmers and farm supervisors, and unskilled blue collar workers insured by the workers' compensation state fund during the study period. The claim rejection rate for farm workers relative to unskilled blue collar workers was 1.69 in 1982 (95% confidence interval [CI] = 1.49–1.92) and 1.47 in 1983 (CI = 1.28–1.68), while the relative rate was 1.06 in 1984 (CI = 0.92–1.22) and

The authors are with the Occupational Medicine Program, University of Washington, Seattle.

Requests for reprints should be sent to Linda Rosenstock, MD, MPH, Occupational Medicine Program, University of Washington, Harborview Medical Center, ZA-66, 325 9th Avenue, Seattle, WA 98104.

This paper was submitted to the journal September 24, 1990, and accepted with revisions April 2, 1991.

TABLE 1—Farm Worker Claims and Estimates of Population at Risk: Washington State, 1982 to 1986

	1982	1983	1984	1985	1986
Claims submitted	6261	6435	6361	5356	5038
rate/1000 person-years	221.7	225.4	206.1	186.4	161.6
Claims accepted	5795	6063	6065	5106	4671
rate/1000 person-years	205.2	212.4	196.5	177.7	149.8
Agricultural workers*	28 238	28 544	30 560	28 738	31 185

*Total hours submitted for billing purposes divided by 2000 for farms, orchards, nurseries, cultivating florists, dairy farms, and livestock and poultry raising.

TABLE 2—Rejection Rates for Farm Workers and Other Occupational Groups: Washington State, 1982 to 1986*

Occupational Group	1982	1983	1984	1985	1986
Farm workers	7.4	5.7	4.6	4.6	7.2
Farmers and farm supervisors	4.7	3.9	3.9	3.5	6.3
Unskilled blue collar workers	4.4	4.0	4.4	4.7	7.5

*All figures are percentages.

similar in succeeding years. Of note, the rejection rate for all groups increased dramatically in 1986.

Discussion

We found a substantial increase in the acceptance of claims filed by farm workers following a judicial ruling that resulted in the removal of a statutory restriction unique to workers in this sector. This increase suggests that the compensation caveat had been a critical factor in rejecting otherwise compensable claims. While this change in the administrative management of farm worker claims is likely the result, and certainly one of the intended benefits, of the Washington State Supreme Court ruling, the limited impact of the ruling on other aspects of workers' compensation is surprising. Unfortunately, although there are indications that coverage of farm workers—in the form of premiums paid by agricultural employers—may have increased slightly after the ruling, utilization of the system does not appear to have increased and indeed claim rates actually decreased in subsequent years.

These trends occurred in spite of the fact that after the ruling the Department of Labor and Industries issued press releases statewide informing all employers then paying premiums. In addition, substantial fines continue to be levied against an employer if a claim is filed and no workers' compensation premiums have been paid. It is possible that farm worker advocates overestimated the effect of the $150 exclusion in inhibiting the filing of claims, but there are many other potential barriers to workers' compensation that were not addressed by the ruling. These include the lack of knowledge of legal rights, of access to health care, of US citizenship, and fluency in the English language.

An inherent limitation of this study is the absence of reliable estimates of the number of farm workers in the state during the study period. Because the number of farm workers may fluctuate from year to year based on weather and market conditions, an independent indicator of the true number of farm workers would have been a useful benchmark against which to measure coverage based on premiums paid. Unfortunately, there is no agency that regularly collects this information. The state OSHA agency (WISHA) does make estimates of the workers in agricultural production on farms with 11 or more employees based on surveys of random samples of employers.[5,6] Using these data as a relative indicator of the true number of farm workers in the state, the ratio of workers' compensation to WISHA estimates was calculated. The ratio was 1.06 in 1982 vs 1.08 in 1986. These figures indicate that there was little difference between the two years in the way employers submitted hours for billing purposes.

If the state is to fulfill its commitment to treat farm workers in the same way as other workers, additional interventions are needed, including informing farm workers directly of their rights and protection under the law. □

Acknowledgments

This paper was presented at the 118th Annual Meeting of the American Public Health Association, October 1990, New York City.

References

1. Coye MJ. The health effects of agricultural production: I. the health of agricultural workers. *J Public Health Policy.* 1985;6:349–370.
2. Strigini P. On the political economy of risk: farm workers, pesticides, and dollars. *Int J Health Serv.* 1983;12:263–292.
3. Moses M. Pesticide-related health problems and farm workers. *Am Assoc Occup Health Nurs J.* 1989;37:115–130.
4. Office of Workers' Compensation Studies; US Department of Labor. *State Workers' Compensation Laws.* Washington, DC: US Department of Labor; January 1990.
5. Division of Industrial Safety and Health. *Washington State Occupational Injury and Illness Survey, 1982.* Olympia, Wash: State of Washington Department of Labor and Industries; 1984.
6. Division of Industrial Safety and Health. *Washington State Occupational Injury and Illness Survey, 1986.* Olympia, Wash: State of Washington Department of Labor and Industries; 1988.

New from NCHS

New from NCHS presents and analyzes statistics from the National Center of Health Statistics that are relevant to public health. The department is edited by Mary Grace Kovar, DrPH. Contributions for this department should be addressed to her as follows: Mary Grace Kovar, DrPH, National Center for Health Statistics, 6525 Belcrest Road, Room 1120, Hyattsville, MD, 20782.

Impairments Due to Injuries

Americans have approximately 19 million physical impairments that are due to injuries. Each year such impairments cause about 275 million days of restricted activity—days spent in bed, absent from work or school, or at less than normal functional capacity. In addition, almost half of these impairments cause a long-term reduction in functional capacity, such as an inability to continue to work or a reduction in the amount or kind of work.

These are some of the key findings from a recent National Center for Health Statistics (NCHS) report, *Selected Impairments Due to Injuries: United States, 1985–87.*[1] The report shows that large numbers of Americans are impaired by injury and that those impairments have a tremendous impact on the lives of individuals and the economy. It confirms recent research about the employment costs of impairments[2] and reinforces the conclusions of a report of the Institute of Medicine that injury is one of the principal public health problems facing America.[3]

The data are from the 1985 to 1987 National Health Interview Survey (NHIS). All national estimates in this report represent annual averages over 3 years (1985 to 1987). The survey has been described[4,5] and the questionnaires have been published.[6] In brief, the NHIS is a continuing national survey of the health of the civilian, noninstitutionalized population of the United States based on an area probability sample. In 1985 to 1987, more than 106 000 households and 276 000 persons were included in this sample after a noninterview rate of 4.3% had been accounted for.

This report presents estimates of the number of impairments, which is larger than the number of people with impairments because a person can have more than one impairment due to injury. A physical impairment is defined as a chronic or permanent defect, usually static in nature, which represents a decrease in or loss of ability to perform various functions. Regardless of the date of onset, impairments due to injuries were categorized as chronic conditions in the NHIS and then were further classified by means of a special supplementary code according to the type of functional impairment and etiology (the classification can be found in the NCHS Medical Coding Manual).[7]

During the mid-1980s, the 236 million people in the civilian, noninstitutionalized population had 71 million impairments. About 26% of these impairments were due to injuries. Of the 19 million impairments caused by injuries, about 12% resulted from injuries that occurred in the 12 months prior to the interview.

Figure 1 shows the types of impairments caused by injuries. Deformities or orthopedic impairments were the most prevalent, accounting for 13.6 million impairments or 71.8% of the impairments due to injuries. The majority of these deformities were of the back and of the lower extremities or parts of the extremities. Hearing impairments were the second most prevalent, adding up to 1.5 million conditions or 8.1% of the total.

This report shows that there are sub-

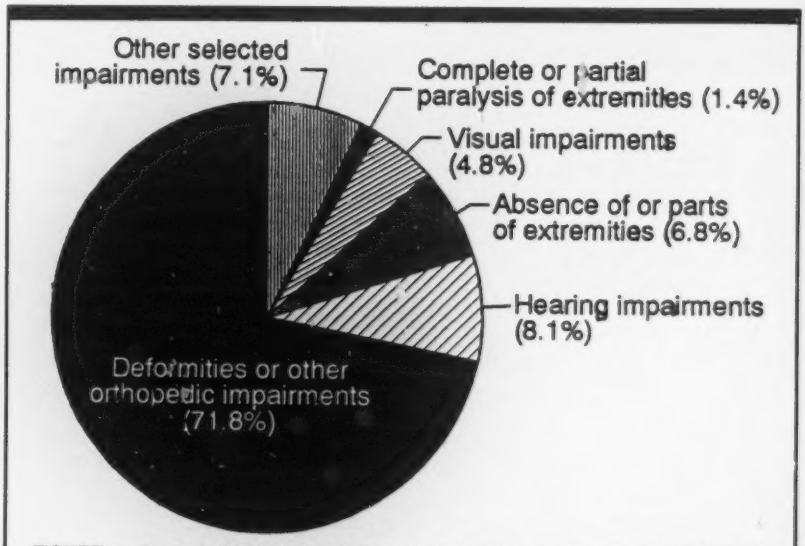

Other selected impairments (7.1%)
Complete or partial paralysis of extremities (1.4%)
Visual impairments (4.8%)
Absence of or parts of extremities (6.8%)
Hearing impairments (8.1%)
Deformities or other orthopedic impairments (71.8%)

FIGURE 1—Percent of selected impairments due to injuries by type of impairment: United States, 1985 to 1987.

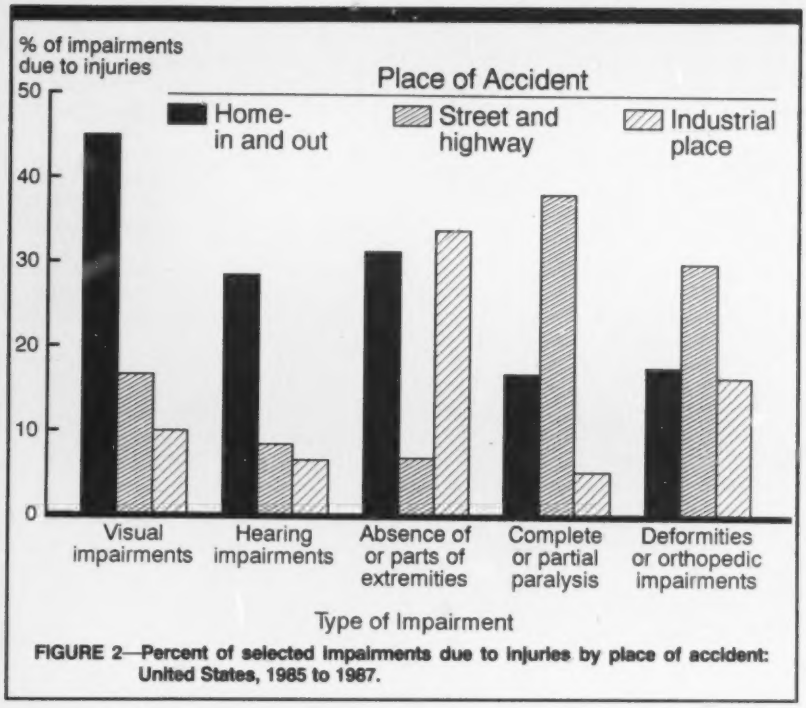


FIGURE 2—Percent of selected impairments due to injuries by place of accident: United States, 1985 to 1987.

stantial differences among various sociodemographic groups. One interesting finding was the difference between men and women. The prevalence rates for all types of selected impairments due to injuries, except for deformities and orthopedic impairments of the back, were higher among men than women.

A second interesting finding was the relationship between injuries and age. Adults aged 45 years and older had higher prevalence rates for all major categories of impairments due to injuries than younger adults. For example, there were 47 deformities or orthopedic impairments per 1000 adults under 45 years of age compared with over 80 per 1000 adults 45 years of age and older.

Another interesting finding was the relationship between injuries and family income. In general, the prevalence of injury-related impairments was inversely associated with income. For example, people in families with incomes of less than $10 000 had higher prevalence rates for visual impairments and deformities or orthopedic impairments than persons with family incomes above $10 000.

The report also provides data on the place where the injury that resulted in the impairment occurred, data that can be used for the targeting of prevention programs. More than one out of four impairments due to injuries occurred in accidents on streets or highways. The second most common place of accident resulting in an impairment was the home: more than one out of five impairments were due to injuries that occurred at home.

Targeting can even be more specific. Some impairments due to injuries were disproportionately represented among primary places of accident (Figure 2). For example, almost one half (45%) of visual impairments and nearly one third (31%) of impairments involving a loss of extremities or parts of extremities were caused by an injury at home. Almost two out of every five cases of paralysis of extremities or parts of extremities due to injuries occurred in accidents on the street or highway. The street or highway was also the most likely place of accident responsible for deformities or orthopedic impairments caused by injuries.

Selected Impairments Due to Injuries shows that the consequences of preventable disability are significant. Many of the impairments due to injuries resulted in a limitation of activity (defined as a long-term reduction in a person's capacity to perform the average kind of activities associated with his or her age group). Almost half (47%) of all physical impairments caused by injuries led to a limitation of activity. For some types of impairments this figure is even higher; for example, 87% of paralysis of extremities or parts of extremities resulted in a limitation of activity.

Injury-related impairments also led to substantial levels of restricted activity (defined as a short-term reduction in a person's activities below his or her normal capacity). 275 million days of restricted activity per year, or about 15 days of restricted activity per impairment, occurred as a result of preventable disabilities. Furthermore, approximately 88 million days per year confining people to their beds because of disability were due to injury-related impairments; this averages to 4.7 bed disability days per impairment.

Physical impairments caused by injuries not only lead to decreases in productive activity, temporary or permanent, but significantly affect overall health status as well. The results of this report point to a strong association between respondent-assessed health status and impairment due to injuries. For all of the major categories of physical impairments, the annual prevalence rates were markedly high among persons reporting a fair or poor health status. For example, the annual prevalence rate for deformities or orthopedic impairments due to injuries among persons reporting a fair or poor health status was almost twice that of persons reporting a good health status (137.6 compared with 71.9 impairments per 1000 persons per year). Thus, an injury-related impairment may, in fact, be a primary cause of differences in self-assessed health status.

Selected Impairments Due to Injuries contains data on the prevalence and associated consequences of preventable disabilities in the United States and also provides information vital to public efforts aimed at prolonging the active life of individuals. With better insight into the population groups most at risk of impairments due to injuries and with better knowledge of the primary places and types of accidents leading to impairments, public programs to reduce one of the major causes of disability in the United States can be more effectively designed and implemented. □

Ann E. Biddlecom, MA
John Gary Collins, MBA

The authors are with the National Center for Health Statistics, Centers for Disease Control, Hyattsville, Md. Requests for reprints should be sent to Ann E. Biddlecom, MA, National Center for Health Statistics, Centers for Disease Control, 6525 Belcrest Road, Hyattsville, MD 20782.

Ordering Information

Copies of the report, *Selected Impairments Due to Injuries: United States, 1985–87* by John

Gary Collins, Vital Health Stat 10(177) can be purchased from the US Government Printing Office, Washington, DC 20402. Order stock no. 017-022-01141-7. The price is $3.25.

For more information about this and other reports or the public-use NHIS data tapes from the National Center for Health Statistics, Centers for Disease Control, write to the Scientific and Technical Information Branch, NCHS, CDC, 6525 Belcrest Road, Hyattsville, MD 20782, or call 301/436-8500.

References

1. Collins JG. *Selected Impairments Due to Injuries: United States, 1985–87.* Washington, DC: US Government Printing Office; 1991. National Center for Health Statistics. *Vital Health Stat.* 10(177).
2. West J, ed. *The Americans with Disabilities Act: From Policy to Practice.* New York: Milbank Memorial Fund; 1991.
3. Institute of Medicine. *Disability in America: Toward a National Agenda for Prevention.* Washington, DC: National Academy Press; 1991.
4. Kovar MG, Poe G. *The National Health Interview Survey Design and Procedures.* Washington, DC: US Government Printing Office; 1985. National Center for Health Statistics. *Vital Health Stat.* 1(18).
5. Massey JT, Moore TF, Parsons VL, Tadros W. *Design and Estimation for the National Health Interview Survey,* 1985–94. Washington, DC: US Government Printing Office, 1989. National Center for Health Statistics. *Vital Health Stat.* 2(110).
6. Dawson DA, Adams PF. *Current Estimates from the National Health Interview Survey: United States, 1986.* Washington, DC: US Government Printing Office, 1987. National Center for Health Statistics. *Vital Health Stat.* 10(164).
7. National Center for Health Statistics. *National Health Interview Survey Coding Manual and Short Index.* Hyattsville, MD: NCHS; 1980.

American Journal of Public Health

January 1992, Vol. 82, No. 1 Established 1911

86389
028
ED KIANKA
CONGRESSIONAL INFORMATION SVC
4520 EAST-WEST HIGHWAY
BETHESDA MD 20814

APHA
American Public Health Association

Featuring Tobacco

Editorials

14 On the Future of Applied Smoking Research: Is It Up in Smoke? *K. E. Bauman*

17 Low Tar, High Toll *K. E. Warner and J. Slade*

19 How Effective Are Excise Tax Increases in Reducing Cigarette Smoking? *J. Wasserman*

Public Health Policy Forum

21 The Future of Public Health *B. Walker, Jr.*

Featuring Tobacco

24 Exposure to Environmental Tobacco Smoke in Naturalistic Settings *K. M. Emmons, D. B. Abrams, R. J. Marshall, R. A. Etzel, T. E. Novotny, B. H. Marcus, and M. E. Kane*

29 Occupational Risks Associated with Cigarette Smoking: A Prospective Study *J. Ryan, C. Zwerling, and E. J. Orav*

33 Misclassification of Smoking Status in the CARDIA Study: A Comparison of Self-report with Serum Cotinine Levels *L. E. Wagenknecht, G. L. Burke, L. L. Perkins, N. J. Haley, and G. D. Friedman*

37 A Prospective Study of the Attributable Risk of Cancer Due to Cigarette Smoking *P.-H. Chyou, A. M. Y. Nomura, and G. N. Stemmermann*

41 Brief Supportive Telephone Outreach as a Recruitment and Intervention Strategy for Smoking Cessation *H. A. Lando, W. L. Hellerstedt, P. L. Pirie, and P. G. McGovern*

47 Adolescent Smoking, Weight Changes, and Binge-Purge Behavior: Associations with Secondary Amenorrhea *J. Johnson and A. H. Whitaker*

55 Predictors of Smoking Prevalence among New York Latino Youth *L. Dusenbury, J. F. Kerner, E. Baker, G. Botvin, S. James-Ortiz, and A. Zauber*

85 Cigarette, Alcohol, and Coffee Consumption and Spontaneous Abortion *B. G. Armstrong, A. D. McDonald, and M. Sloan*

87 Cigarette, Alcohol, and Coffee Consumption and Prematurity *A. D. McDonald, B. G. Armstrong, and M. Sloan*

91 Cigarette, Alcohol, and Coffee Consumption and Congenital Defects *A. D. McDonald, B. G. Armstrong, and M. Sloan*

94 The Effect of State Cigarette Tax Increases on Cigarette Sales, 1955 to 1988 *D. E. Peterson, S. L. Zeger, P. L. Remington, and H. A. Anderson*

96 Mass Media–Led Antismoking Campaign Can Remove the Education Gap in Quitting Behavior *P. Macaskill, J. P. Pierce, J. M. Simpson, and D. M. Lyle*

99 Demographic and Socioeconomic Differences in Beliefs about the Health Effects of Smoking *R. C. Brownson, J. Jackson-Thompson, J. C. Wilkerson, J. R. Davis, N. W. Owens, and E. B. Fisher*

(continued page 1)

Contents

American Journal of Public Health

"Information for Authors" appears on page 138 of this issue and may also be obtained from the Journal Editorial Office.

American Journal of Public Health
1015 Fifteenth Street, NW
Washington, DC 20005

APHA
American Public Health Association

January 1992, Vol. 82, No. 1 ISSN: 0090-0036 Established 1911

103 Geographic and Gender Variations in Total Tobacco Use *D. R. Shopland, S. J. Niemcryk, and K. M. Marconi*

107 Handmade Cigarettes: It's the Tube That Counts *M. J. Kaiserman and W. S. Rickert*

Articles

59 Variations in Asthma Hospitalizations and Deaths in New York City *W. Carr, L. Zeitel, and K. Weiss*

66 Diversity of Tick Species Biting Humans in an Emerging Area for Lyme Disease *R. P. Smith, Jr., E. H. Lacombe, P. W. Rand, and R. Dearborn*

70 Prevalence of Human T-Lymphotropic Virus in Civilian Applicants for the United States Armed Forces *C. R. Roberts, D. R. Fipps, J. F. Brundage, S. E. Wright, M. Goldenbaum, S. S. Alexander, and D. S. Burke*

74 Correlates of Using Mental Health Services: Implications of Using Alternative Definitions *D. Mechanic, R. J. Angel, and L. Davies*

79 Evaluation of a Self-help Dietary Intervention in a Primary Care Setting *S. A. A. Beresford, E. M. Z. Farmer, L. Feingold, K. L. Graves, S. K. Sumner, and R. M. Baker*

Public Health Briefs

109 Improvement in Intergenerational Health *J. F. Fries, C. A. Williams, and D. Morfeld*

112 The Effect of Legal Drinking Age on Fatal Injuries of Adolescents and Young Adults *N. E. Jones, C. F. Pieper, and L. S. Robertson*

115 Evaluating the Treatment of Sexually Transmitted Diseases at an Urban Public Hospital Outpatient Clinic *P. G. Shekelle and J. Kosecoff*

117 Surveillance of Sentinel Occupational Mortality in the District of Columbia: 1980 to 1987 *A. Cottrell, E. Schwartz, R. Sokas, V. Kofie, and L. Welch*

120 Maternal Recall Error of Child Vaccination Status in a Developing Nation *J. J. Valadez and L. H. Weld*

Other Journal Departments

2 Masthead

5 Advertisers' Index

123 Letters to the Editor

129 Book Corner

133 Nomination Forms

138 Information for Authors

145 Job Opportunities

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.

APHA OFFICERS

Joyce C. Lashof, MD, *President*

C. William Keck, MD, MPH, *Immediate Past President*

Helen Rodriguez-Trias, MD, *President-Elect*

Arnold E. Greenberg, SM, *Vice-President (USA)*

Trevor Hancock, MB, MHSc, *Vice-President (Canada)*

Julio Teja Perez, MD, *Vice-President (Latin America)*

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Frances M. Veverka, RS, MPH, *Chair, Executive Board*

George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*

William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD

Frances M. Veverka, RS, MPH (1992), *Chair*

Marilyn Aguirre-Molina, EdD, MS (1993)

Brian A. Dobrow, MPH (1992)

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Kristine M. Gebbie, BSN, MN (1994), *Vice-Chair*

Alan R. Hinman, MD, MPH (1995)

Janet Heinrich, DrPH, MPH (1993)

Margaret H. Jordan, RN, MPH (1992)

C. William Keck, MD, MPH, *Immediate Past President*

Margo D. Kerrigan, MPH (1995)

Joyce C. Lashof, MD, *President*

Barry S. Levy, MD, MPH (1993)

Helen Rodriguez-Trias, MD, *President-Elect*

Allan G. Rosenfield, MD (1994)

Reed V. Tuckson, MD (1994)

Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO

Willam H. McBeath, MD, MPH, *Executive Director*

Joanne E. Lukomnik, MD, *Chair, Action Board*

George Pickett, MD, MPH, *Chair, Program Development Board*

George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc., 1015 Fifteenth St., NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50,000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health,* 1015 Fifteenth St., NW, Washington, DC 20005.

Information on Manuscript Preparation is published in this Journal in January. The Journal publishes reports of original research, demonstrations, evaluations, and other articles covering current aspects of public health. Manuscripts are accepted for consideration with the understanding that they, or their essential substance, have not been published previously or accepted for publication and are not under consideration elsewhere.

Subscription price, individual, is $80/year US, $120/year international; **institutional,** $160/year US, $240/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to American Journal of Public Health, 1015 Fifteenth St., NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact: Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal of APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact: Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols. 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS Bibliographic Retrieval Service Inc, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc., Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Exposure to Environmental Tobacco Smoke in Naturalistic Settings

ABSTRACT

Background. Exposure to environmental tobacco smoke (ETS) has been identified as a risk factor for chronic disease among nonsmokers. Results of epidemiological surveys suggest that the majority of nonsmokers have regular ETS exposure. However, little is known about the topography of exposure.

Methods. An exposure diary was used by 186 nonsmokers to self-monitor ETS exposure over a 7-day period. Subjects also completed a questionnaire that assessed their patterns of ETS exposure.

Results. The primary source of ETS exposure was the workplace, except when there was a smoker in the household, in which case the household was the primary source. The presence of a smoker in the household resulted in higher levels of exposure both at work and in other locations when compared with subjects without household exposure. Subjects' assessments of exposure on the questionnaire were consistently lower than their self-monitored levels. This finding suggests that general exposure ratings underestimate exposure.

Conclusions. This study provides a new understanding of the patterns of ETS exposure and may help guide the development of policies and interventions designed to reduce ETS exposure. (*Am J Public Health.* 1992;82:24–28)

Karen M. Emmons, PhD, David B. Abrams, PhD, Robert J. Marshall, PhD, Ruth A. Etzel, MD, PhD, Thomas E. Novotny, MD, Bess H. Marcus, PhD, and Margaret E. Kane, BS

Introduction

Exposure to environmental tobacco smoke (ETS) has been identified as a risk factor for lung cancer and chronic obstructive pulmonary disease among nonsmokers.[1–7] Exposure to ETS increases acute respiratory symptoms and leads to a decrement in pulmonary function.[8–12] Results of population-based surveys suggest that 88% of nonsmokers are aware of the negative health consequences of passive exposure to tobacco smoke.[13] Despite this general awareness, however, exposure to ETS is pervasive. More than 55% of the respondents to the Adult Use of Tobacco Survey reported that their workplace did not have any smoking restrictions.[13] In an interview study, Cummings and colleagues found that over 75% of nonsmokers (n = 663) reported ETS exposure in the 4 days prior to their participation in the study.[14]

We must assess the pattern and degree of ETS exposure if we are to understand the disease risk associated with exposure and evaluate the success of interventions designed to minimize such exposure. The focus of this paper is on the measurement of acute daily exposure in field settings. One problem with measuring ETS exposure is that estimates of exposure are often based on global survey measures that rely on subjective estimates of quantity (e.g., "a little," "some," or "a lot"),[15] often over nonuniform periods of time. In addition, survey measures have been incomplete in assessing the parameters of exposure, such as location, source, intensity, duration, and ventilation.[16–18] Therefore, although some global measures have been useful, it is difficult to obtain specific estimates of the parameters of exposure on the basis of information obtained with survey measures.[18]

Little is known about the precise parameters of exposure among nonsmokers in the natural environment. Knowledge about these parameters of exposure is important for intervention, public health education, and policy efforts designed to reduce overall ETS exposure.[7,19] In this paper, by estimating daily exposure, we provide situation-specific information (place, time, intensity, and proximity) that may be useful for individuals in planning avoidance or reduction strategies to minimize their exposure to ETS and for policymakers to improve the protection of the public's health.[20]

The data presented here are from a larger study designed to develop interventions to reduce ETS exposure in the workplace. The purpose of the present study was to develop a method for self-monitoring ETS exposure continuously over a 7-day period, to use this method to document the frequency of exposure in natural

Karen M. Emmons, David B. Abrams, and Bess H. Marcus are with Brown University and The Miriam Hospital, Division of Behavioral Medicine, Providence, RI. Robert J. Marshall is with the Office of Health Promotion, The Rhode Island Department of Health. Ruth A. Etzel is with the Center for Environmental Health and Injury Control, Centers for Disease Control, Atlanta, Ga, and Thomas E. Novotny is with the Office on Smoking and Health, Centers for Disease Control, Rockville, Md. Margaret E. Kane is with the Rhode Island Lung Association, Providence, RI.

Address requests for reprints to Karen M. Emmons, PhD, Division of Behavioral Medicine, The Miriam Hospital—RISE Building, 164 Summit Avenue, Providence, RI 02906.

This paper was submitted to the journal August 9, 1990, and accepted with revisions May 6, 1991.

settings, and to document selected exposure parameters. Results obtained with this self-monitoring technique were also compared with global ratings typically used in survey research.

Methods

Subjects

Recruitment efforts were focused on nonsmokers in workplace settings that were selected to have a wide range of exposure to ambient ETS. The work sites ranged from those with minimal restriction of smoking and high levels of exposure (e.g., long-term care and psychiatric facilities, chemical dependency treatment centers, a VA Hospital) to work sites with extensive smoking restrictions and low exposure (e.g., the state health department and community hospitals). Approximately two thirds of the participants were recruited directly from workplaces, and the remainder were recruited via newspaper advertisements.

Of 224 nonsmokers who volunteered to participate in a study of passive smoking, 38 were excluded because of incomplete data. The demographic characteristics of the sample are presented in the Table. The subjects' mean age was 41.7 years (SD = 12.2). Their mean educational level was 15.8 years (SD = 2.8), and their income range (modal) was $30 000 to $49 999. Ninety percent of the subjects worked outside the home. Eighty-four percent of those who worked outside the home (75.6% of the total sample) reported being regularly exposed to smoking in the workplace. Thirty percent of the subjects reported living with one or more (range: one to four) smokers.

Measures

Exposure questionnaire. The exposure questionnaire provided a global estimate of subjects' exposure to ETS. The questionnaire was designed to elicit information about the respondents' perception of their level of ETS exposure (number of minutes per day, intensity of exposure) at various locations (work, home, and other). Several knowledge and attitude questions from the Office on Smoking and Health's 1988 Adult Use of Tobacco Survey[13] were included; results based on these items will be presented elsewhere.

Exposure diary. The exposure diary consisted of small (3.5 × 5 inches) printed cards (see Figure 1). Each entry required only a simple check mark or short response, which subjects were asked to make at the time of exposure. Subjects were instructed to define "exposure" as contact with an individual who was actively smoking. The parameters assessed by the diary included number and duration of exposures location of exposure (work, home, other), intensity of exposure (number of smokers), and distance from the source of exposure ("near" = less than 5 feet; "far" = more than 5 feet).

Procedure

A research assistant met with all subjects on two separate occasions (1 week apart) to administer the exposure questionnaire, to train subjects in the use of the exposure diary, and to collect the data. To facilitate appropriate use of the exposure categories, the research assistant had the subjects code several sample exposure scenarios and provided feedback about their coding accuracy. Self-monitoring began on the day after the training session and continued for 7 consecutive days. The distribution of the initial day of monitoring across the days of the work week was approximately equal.

Ideally, the global exposure questionnaire and the exposure diary would have been collected over the same time interval. However, the act of self-monitoring (keeping the diary) would presumably have enhanced the accuracy of the global responses and confounded any meaningful comparison of the two approaches. To minimize this confound, the exposure questionnaire was administered first, asking about the preceding 7-day period. After they completed the questionnaire, the subjects were trained in the use of the 7-day exposure diary. To make the two 7-day time periods as comparable as possible, the subjects were asked to participate in the study during a 2-week time period that reflected their typical daily lifestyle (e.g., at a time when they were not on vacation).

Results

Data Stability

Exposure duration was similar when examined across different segments of the monitoring period (7 days vs last 3, 2, and 1 days); there were no significant differences among the four sampling intervals, and the Pearson product-moment correlations among the intervals ranged from .85 to .97. Therefore, average daily exposure based on information in the 7-day diary was used for all analyses.

TABLE 1—Description of Study Sample (n = 186)

	n	%
Sex		
Male	69	37
Female	117	63
Marital status		
Single/divorced	58	31
Married/cohabiting	128	69
Ethnic group		
White	181	97
Black	5	3
Smoking status		
Never a smoker	115	62
Ex-smoker	71	38

Data Reduction

Individual exposure durations were calculated from the start and stop times of each exposure noted in the diary. Total duration of exposure (in minutes per day) was derived from the sum of the individual episodes of exposure. Exposure duration was computed for the entire 7-day monitoring period. To assess the potential influence of reactivity and response burden, the relative stability of shorter sampling intervals (24, 48, and 72 hours vs 7 days) was also computed.

Parameters of Exposure Determined by the Exposure Diary

Duration of exposure by location. In the total sample, as indicated in Figure 2, significantly more of the exposure was encountered at work, with home and "other" contributing less to overall exposure. Approximately 50% of the exposure was in the workplace; the home environment accounted for 10% of the total daily exposure. However, when those with and those without a smoker in the household were examined separately, we found that subjects who lived with a smoker received more exposure in the home than in the workplace, omnibus $F(33) = 4.47$, $P \leq .01$ (see Figure 3).

The presence of smokers in the home had a significant effect on duration of exposure across specific settings. Subjects without smokers at home reported that the majority of their exposure was in the workplace (mean = 36.1 minutes per day, 95% confidence interval [CI] = 22.7–49.5), with very little either at home (mean = 1.4 minutes per day, 95% CI = 0.05–2.75) or in other locations (mean = 13.1 minutes per day, 95% CI = 8.75–17.4). However, subjects who lived with smokers had virtually equivalent exposure across all three settings (work: mean = 29.4 minutes per day,

EXPOSURE DIARY SIDE 1

Date __/__/__ Time Awake: __:__ Time Asleep __:__

Time Start	Time Stop	Place	# of Smokers	Distance
		W H O		N F
		W H O		N F
		W H O		N F
		W H O		N F
		W H O		N F
		W H O		N F
		W H O		N F
		W H O		N F
		W H O		N F

EXPOSURE DIARY SIDE 2

Time start: time that you 1st encounter cigarette smoke

Time stop: time that exposure to smoke ends

Place: W = work H=home O=other circle the letter that most closely describes your location at the time you encounter cigarette smoke

Number of Smokers: the number of people that are smoking

Distance: "N" (Near): less than 5 feet away "F" (Far): more than 5 feet away

FIGURE 1—Exposure diary.

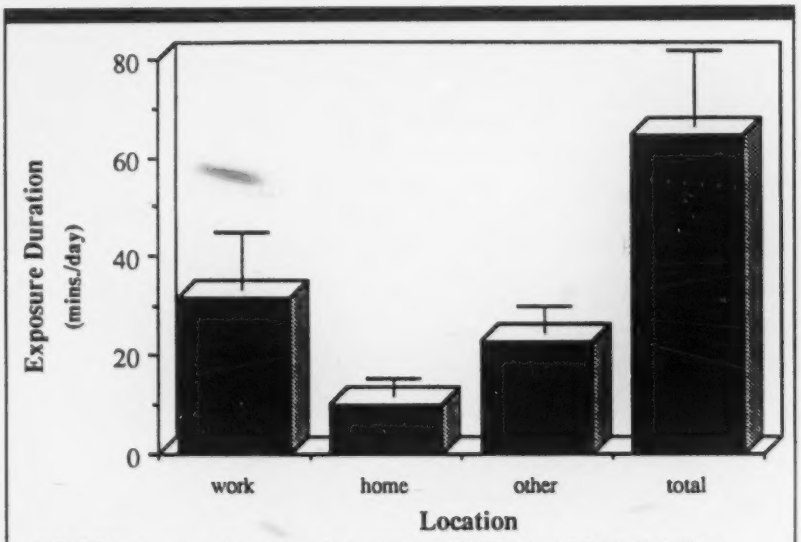


FIGURE 2—Amount of exposure to environmental tobacco smoke by location.

95% CI = 7.01–51.80; home: mean = 31.2 minutes per day, 95% CI = 21.6–40.8; other: mean = 27.1 minutes per day, 95% CI = 15.1–39.1).

To further examine differences between subjects with differing levels of exposure, we used a tertile split based on the average daily exposure to divide the sample into three levels of exposure: low (0–7 minutes per day), moderate (8–44 minutes per day), and high (≥45 minutes per day). These categories are somewhat arbitrary, because no standard has been set in the field about the severity of various gradients of exposure. Subjects in the high-exposure category were more likely to have smokers both at home, $F(2) = 4.93$, $P \leq .01$, and at work, $F(2) = 3.11$, $P \leq .05$. They also had more friends outside the workplace who were smokers, $F(2) = 5.31$, $P \leq .01$.

Intensity of exposure. Ratings of distance from the source of exposure were used as a measure of exposure intensity. We compared the percentages of exposure at each location (home, work, other) that occurred near and far from the subject (see Figure 4). Significantly more of the exposure occurring at home was close to the subject, and therefore more intense, compared with exposure at work and other locations.

Relationship between the Exposure Diary and Exposure Questionnaire

Subjects were asked on the exposure questionnaire to provide a global rating of their ETS exposure on a three-point scale (low, moderate, high). Fifty-three percent of the sample rated their exposure as low. Thirty-six percent rated their exposure as moderate, and 11% as high. Subjects who rated their overall exposure as low had an average of 18.59 minutes of exposure per day (SD = 27.08, range = 0.0–203.0), compared with an average of 110.28 minutes per day (SD = 142.82, range = 0.0–603) for subjects who rated their exposure as moderate and an average of 132.24 minutes per day (SD = 115.15, range = 8.67–408.0) for those who rated their exposure as high.

Correlational analyses were used to compare the exposure diary data with global estimates from the exposure questionnaire across locations (home, work, other). The correlation was strongest between global and diary estimates of exposure at home, $r(185) = .62$, $P \leq .001$ (95% CI = .58–.87), but all correlations (work, $r = .40$; other, $r = .29$; overall, $r = .44$) were significant at $P \leq .01$.

Relationships between Demographics and Exposure

Exposure was significantly related to education, $r(185) = -.22$, $P \leq .01$ (95% CI = −.37 to −.08): exposure was higher among less educated subjects. The relationship between gender and exposure was marginally significant ($P = .10$), indicating a trend for men to have higher levels of exposure than women. No other relationships between demographics and exposure were significant.

Discussion

This study represents one of the first attempts to develop a prospective, diary measure of daily exposure to ETS, to examine sources of exposure for a 1-week period, and to compare the diary with a global questionnaire measure of exposure. We found the exposure diary to be a relatively simple and effective method of assessing ETS. Furthermore, the data were quite stable, as evidenced by the high correlation between the 1-, 2-, 3-, and 7-day blocks of self-monitored exposure. Thus, a 1- or 3-day diary may be preferred to reduce response burden.

The majority of ETS exposure occurred in the workplace. The percentage of subjects reporting exposure at work (75.6%) replicates that found by Cummings et al.[14] (75%) in a different sample. Furthermore, both the duration and the intensity of exposure were greater for subjects with a smoker in the household,

regardless of the degree of exposure in the workplace. Importantly, subjects with smokers in the home had virtually equivalent exposure across all settings, whereas subjects who did not live with smokers had the majority of their exposure at work and very little in other settings. This suggests that nonsmokers who live with smokers are exposed outside the home to a greater degree than nonsmokers who have no smokers in the household. The reasons for this are unclear, but it is possible that these nonsmokers have less motivation to avoid ETS exposure in general, because they cannot avoid it at home. As expected, subjects in the highest exposure category had regular ETS exposure both at work and at home. These data suggest that for intervention efforts to be effective for those subjects at highest risk, they need to be focused on all sources of exposure—particularly if there is a smoker in the household.

The subjects' own assessments of their exposure levels on the questionnaire were consistently lower than the levels determined by the diary, suggesting that studies utilizing global exposure ratings from surveys may underestimate actual exposure. This may be due in part to the insidious nature of exposure—a high level of vigilance is necessary if accurate estimates are to be obtained. In this case, self-monitoring techniques may increase the accuracy of exposure estimates. Survey-based estimates may also be lower because of the unpredictability of ETS exposure. As noted earlier, the exposure diary and the exposure questionnaire relate to different time periods, 1 week apart. This was necessary to minimize the confound that might have resulted from the increased awareness of exposure due to self-monitoring, which might have influenced responses to the exposure questionnaire. Since both time periods were regarded as typical, and since there was high stability between the 1-, 2-, 3-, and 7-day average exposure durations, it is unlikely that the different sampling intervals alone accounted for the underestimates of the exposure questionnaire.

The subjects in this study were volunteers who responded to advertisements for a study on passive smoking, and thus were a motivated sample of convenience. A selection bias or social desirability bias may have been operating, so that participants may have been more vigilant about noting their exposures to ETS than persons in a random sample might have been. However, the recruitment protocol was designed to target individuals with a wide range of ETS exposure in order to increase the representativeness of the sample. The amount of exposure across individuals in the sample varied considerably, suggesting that the population was not homogeneous. In addition, the demographics of this sample were virtually identical to those found in population-based samples by, for example, Marshall et al.[21] Thus, it does not appear that a significant selection bias was operating.

One concern about the use of the diary is the recall bias that may be introduced if exposures are not recorded at the time they occur. Although some diary entries may have been made after the actual exposure, every effort was made to minimize such occurrences (training emphasized immediate completion, and the diary was designed to minimize response burden). In addition, the subjects were asked during debriefing about the case of diary completion, and few reported having difficulty completing the diary as instructed.

This study provides further understanding of the patterns of ETS exposure in natural settings. It also sheds light on the relationship between retrospective global estimates of exposure and prospective recording of actual exposure as it oc-

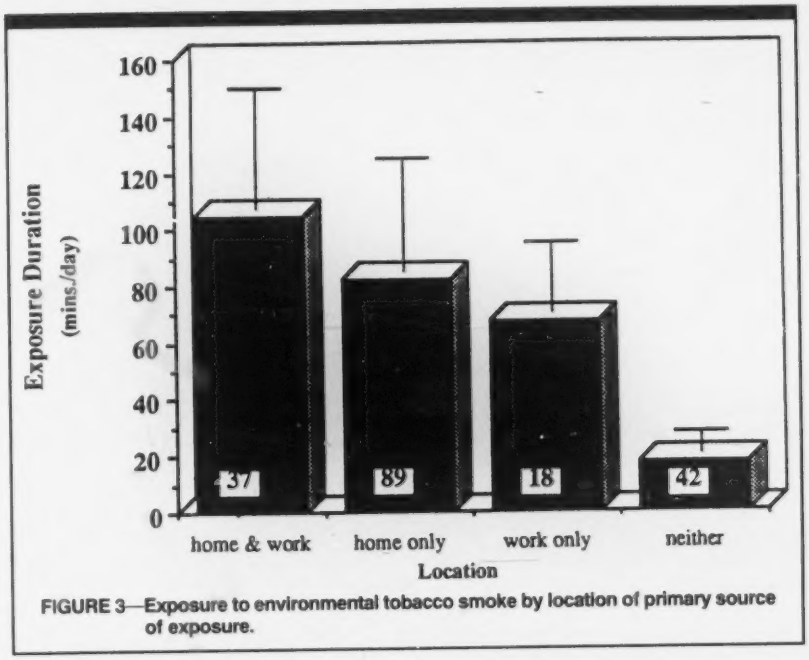


FIGURE 3—Exposure to environmental tobacco smoke by location of primary source of exposure.

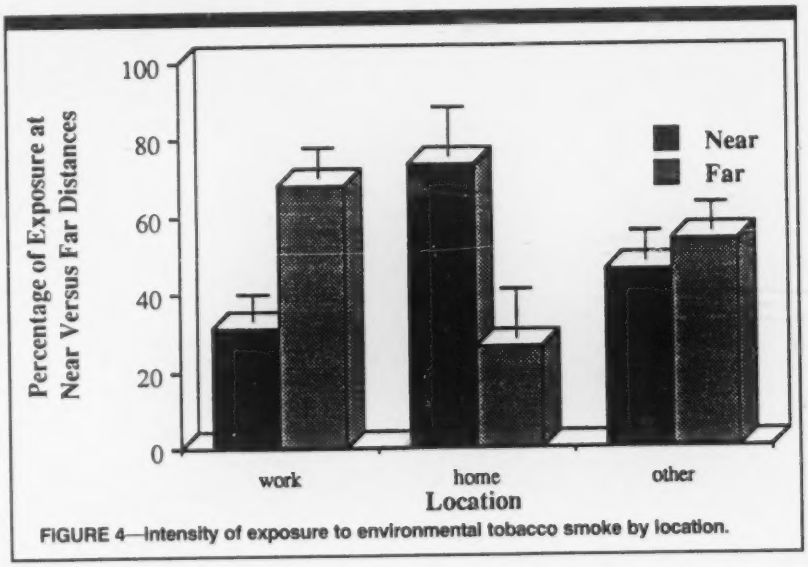


FIGURE 4—Intensity of exposure to environmental tobacco smoke by location.

curs in the field. This comparison is tempered by the methodological limitations that precluded assessment of both measures in the same time period. These data may be useful in developing effective public health policies and in designing intervention programs targeted at both individuals and organizational settings to reduce nonsmokers' exposure to ETS. □

Acknowledgments

This research was supported in part by a cooperative agreement from the Centers for Disease Control (450-CCU-102970) and the National Cancer Institute (1-RO3-CA-48437 and CA 38309). An earlier version of this paper was presented at the 10th annual meeting of the Society of Behavioral Medicine, April 1, 1989, San Francisco, Calif. The authors wish to thank Sheilarae Carpentier, Elaine Taylor, Suzanne Sales, Cathy Draycott, Nancy Keenan, and Alice Baker for their efforts on this project.

References

1. Correa P, Pickle LW, Fontham E, Lin Y, Haenszel W. Passive smoking and lung cancer. *Lancet.* 1983;2:595–597.
2. Doll R, Peto R. Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and lifelong nonsmokers. *J Epidemiol Community Health.* 1978;32:303–313.
3. Hirayama T. Non-smoking wives of heavy smokers have a higher risk of lung cancer: a study from Japan. *Br Med.* 1981;282:183–185.
4. Hirayama, T. Passive smoking and lung cancer. *Br Med.* 1981;282:1393–1394.
5. Hirayama, T. Passive smoking and lung cancer: consistency of association. *Lancet.* 1983;2:1425–1426.
6. Miller GH, Gerstein DR. Life expectancy of non-smoking men and women. *Public Health Rep.* 1983;98:343–349.
7. US Department of Health and Human Services. *The Health Consequences of Involuntary Smoking: A Report of the Surgeon General.* Washington, DC: US Govt Printing Office; 1986. DHHS publication CDC 87–8398.
8. Aronow WS. Effect of passive smoking on angina pectoris. *N Engl J Med.* 1978;229: 21–24.
9. Dahms TE, Bolin JF, Salvin RG. Passive smoking: effects on bronchial asthma. *Chest.* 1981;80:530–534.
10. Kauffmann F, Tessler JF, Oriol P. Adult passive smoking in the home environment: a risk factor for chronic airflow limitation. *Am J Epidemiol.* 1983;117:269–280.
11. US Department of Health and Human Services. *The Health Consequences of Smoking—Chronic Obstructive Lung Disease: A Report of the Surgeon General.* Washington, DC: US Govt Printing Office; 1984. DHHS publication PHS 84–50205.
12. White JR, Froch HF. Small-airways dysfunction in non-smokers chronically exposed to tobacco smoke. *N Engl J Med.* 1980;302:720–723.
13. Centers for Disease Control. *US Department of Health and Human Services Report of the 1986 Adult Use of Tobacco Survey.* Rockville, Md: US Department of Health and Human Services, Public Health Service; 1989.
14. Cummings KM, Markello SJ, Mahoney M, Bhargava AK, McElroy PD, Marshall JR. Measurement of current exposure to environmental tobacco smoke. *Arch Environ Health.* 1990;45:74–79.
15. Jarvis MJ, Pesoe HT, Feyerabend C, Vesey L, Saloojee Y. Biochemical markers of smoke absorption and self-reported exposure to passive smoking. *J Epidemiol Community Health.* 1984;38:335–339.
16. National Research Council. *Environmental Tobacco Smoke: Measuring Exposure and Assessing Health Effects.* Washington, DC: National Academy Press; 1986.
17. Coultas DB, Howard CA, Peake GT, et al. Salivary cotinine levels and involuntary tobacco smoke exposure in children and adults in New Mexico. *Am Rev Respir Dis.* 1987;136:305–309.
18. Jarvis MJ, Russell MAH, Feyerabend C. Absorption of nicotine and carbon monoxide from passive smoking under natural conditions of exposure. *Thorax.* 1983;38:829–833.
19. US Department of Health and Human Services. *Reducing the Health Consequences of Smoking—Twenty-Five Years of Progress: A Report of the Surgeon General.* Washington, DC: US Govt Printing Office; 1989. DHHS publication CDC 89–8411.
20. Abrams DB, Emmons KM, Niaura RS, Goldstein MG, Sherman C. Tobacco dependence: integrating individual and public health perspectives. In: Nathan PE, Langenbucher JW, McCrady BS, Frankenstein W, *The Annual Review of Addictions Treatment and Research.* Vol 1. In press.
21. Marshall RJ, Buechner J, McCombs S, Cataldo J. The prevalence of household passive smoke exposure: a summary of population based data. Presented at the meeting of the Society of Behavioral Medicine; 1989; San Francisco, Calif.

Occupational Risks Associated with Cigarette Smoking: A Prospective Study

James Ryan, MD, MPH, Craig Zwerling, MD, PhD, MPH, and Endel John Orav, PhD

ABSTRACT

Background. Studies have indicated that cigarette smokers have more occupational accidents and injuries and use more sick time and health benefits than nonsmokers, thereby producing sizeable costs for employers. However, they usually have not controlled for other possible sources of these costs. We analyzed occupational costs associated with smoking while adjusting for a number of potential confounders.

Methods. We conducted a prospective, controlled study of the association between smoking and employment outcomes in 2537 postal employees, adjusting for age, gender, race, drug use, job category, and exercise habits.

Results. For smokers, the relative risk for turnover was 1.01 (95% confidence interval [CI], 0.83–1.21); for accidents 1.29 (CI, 1.07–1.55); for injuries 1.40 (CI, 1.11–1.77); for discipline 1.55 (CI, 1.19–2.02). Their mean absence rate was 5.43% compared with 4.06% for nonsmokers.

Conclusions. Our study shows that cigarette smoking is associated with adverse employment outcomes after controlling for a number of possible confounders. This finding has implications for companies formulating smoking policies and considering the establishment of smoking cessation programs. (*Am J Public Health.* 1992;82:29–32)

Introduction

The social costs of cigarette smoking, including occupational risks, have been widely discussed since the original surgeon general's report on smoking and health.[1,2] A 1986 study of public employees found that cigarette smokers use 23% more sick leave than nonsmokers.[3] Studies of student nurses in 1983 and 1988 found smokers absent 77% more frequently than nonsmokers[4] and a significant relationship ($P < .008$) between number of cigarettes smoked and days absent.[5] Smokers have been reported to use more health benefits[6,7] and have more occupational accidents and injuries. A 1966 study of machinists reported an injury rate of 18% for smokers compared with 10% for nonsmokers; 26% of injured smokers reported multiple injuries vs 13% of injured nonsmokers.[8] A 1973 study involving ceramics workers[9] found that 57% of smokers sustained occupational injuries and 20% lost time from work as a result of injury. Among nonsmokers, only 28% sustained injuries and 10% lost time. These associations account for sizable costs for employers of smokers.[10,11]

Smoking rates vary with demographic variables, including age, sex, and race.[2] Smoking is also associated with use of drugs[12] and excessive alcohol.[13,14] Since such factors may be associated with adverse health outcomes, studies of risks associated with smoking should control for these potential confounding variables.

We present a prospective examination of the association between cigarette smoking and rates of employee turnover, absenteeism, accidents, injuries, and discipline while controlling for age, gender, race, job category, exercise habits, and drug abuse. It represents a secondary analysis of data collected in a previous study of the predictive value of preemployment drug testing.[15]

Methods

Cohort Definition

All Boston postal service applicants for permanent nonprofessional positions examined between September 24, 1986, and January 6, 1989, were potential study subjects. Candidates for temporary and professional positions were excluded, the latter because of their small number (n = 11) and different socioeconomic and educational backgrounds.

From September 24, 1986, to January 6, 1989, 4957 applicants received preemployment examinations. We excluded 56 applicants because data-entry errors made it impossible to track their outcome variables, 60 Hispanics (of whom 24 were hired) and 2 Native Americans (1 hired) because the numbers were too small for statistical analysis, and 42 applicants with a history of treatment for drug abuse and

James Ryan is with the East Boston Neighborhood Health Center, and the Occupational Medicine Program, Department of Environmental Health, Harvard School of Public Health, Boston, and the Department of Community Medicine, Tufts University School of Medicine, Boston; Craig Zwerling is with the Institute of Agricultural Medicine and Occupational Health, Department of Preventive Medicine and Environmental Health, College of Medicine, University of Iowa; Endel John Orav is with the Department of Biostatistics, Harvard School of Public Health, Boston.

Requests for reprints should be sent to James Ryan, MD, PO Box 341, Jamaica Plain, MA 02130-0341.

This paper was submitted to the journal March 4, 1991, and accepted with revisions August 22, 1991.

7 who tested positive for opiates under a prior agreement with postal service management to fully evaluate such persons. Thus, the study population was not entirely representative of all postal service employees hired during the study period. Persons removed from the study were excluded without consideration of their smoking status. Of the remaining 4797 applicants, 2 for whom data on smoking status were unavailable were excluded, leaving 4795 potential subjects for this study; 2537 applicants hired before January 28, 1989, were included. Both mean and median duration of follow-up were 406 days, with a range of 28 to 896 days.

Smoking Status

Smoking status was categorized as smoker or nonsmoker using self-reported data. Persons who had smoked three or more cigarettes per day over the previous month were classified as smokers. All other subjects were classified as nonsmokers. One tenth of 1% of subjects were smokers who smoked less than three cigarettes per day. Any misclassification bias introduced by this classification system would therefore be negligible and, in any case, toward the null.

Self-reported information on smoking status was used because of its convenience, simplicity, economy, and reliability. Self-reported data have been demonstrated to have sensitivity and specificity in excess of 90% when compared with objective measurements of saliva cotinine,[16] serum thiocyanate,[17] and expired carbon monoxide.[17,18]

Confounding Factors

A number of confounding factors might be associated with both the risk variable (cigarette smoking) and the outcome variables (turnover, absenteeism, accidents, injuries, and discipline). We therefore collected data on age, gender, race, job classification, drug use, and exercise habits.

Race was classified by the examiner as White, Black, or Asian. Job classifications included letter carrier, special delivery messenger, motor vehicle operator, postal clerk, letter sorting machine clerk, computer forwarding system clerk, mail handler, and maintenance worker, divided into two or three categories according to endpoint. Applicants were divided into "aerobic exercisers" (strenuous exercise for 20 minutes, three times a week), "regular exercisers," and "nonexercisers" based on self-report. A urine drug screen was performed for marijuana, cocaine, amphetamines, barbiturates, benzodiazepines, methadone, methaqualone, opiates, phencyclidine, and propoxyphene as reported elsewhere.[15] In the analysis, drug test results were classified into four categories: negative for all drugs, positive for marijuana only, positive for cocaine only, and positive for other drugs or multiple drugs.

Endpoints

Five endpoints were considered in the analysis: absence rate, time to termination, time to first accident, time to first injury, and time to first discipline.

Absence rates. The absence variable was a function of sick leave (SL), leave without pay (LWOP), and work hours calculated according to the following formula:

$$\text{Absence rate} = \frac{\text{SL hours} + \text{LWOP hours}}{\text{SL hours} + \text{LWOP hours} + \text{work hours}}$$

Time to termination. Individuals whose employment with the postal service had terminated were defined as having turned over. Termination was subdivided into two categories, voluntary (death, resignation, or transfer to another federal agency) and involuntary (fired).

Time to first accident. Records of work-related accidents revealed a relatively low rate of multiple accidents. Of 502 employees reporting accidents, 76 had two accidents and 14 had three. We therefore analyzed time to first accident rather than accident rates.

Time to first injury. We found a relatively low rate of multiple injuries. Of 312 employees who reported occupational injuries, 41 had two injuries and 8 had three. We therefore analyzed time to first injury rather than injury rates.

Time to first report of disciplinary action. The postal service labor agreements provide for progressive levels of discipline up to removal. The time to first discipline was calculated from labor relations department records. All endpoints were followed through March 11, 1989.

Statistics

We used the Cox proportional hazards model[19] from the SAS statistical package's SUGI supplement[20] with nonsmokers as the baseline group to assess whether cigarette smoking was associated with turnover, accidents, injuries, or discipline. Thus, a significantly positive Cox regression coefficient for smokers indicates a significantly increased risk for that endpoint and a negative Cox regression coefficient indicates a decreased risk.

Because 25% of the subjects had a 0% absence rate, it was not possible to use linear regression with absence rate as the outcome variable. Subjects who were absent were divided into approximately equal quartiles according to their absence rates: 0% absence rate (n = 646), 0% to 1.52% (n = 478), 1.52% to 3.04% (n = 468), 3.04% to 5.60% (n = 471), and absence rate > 5.60% (n = 474). The SAS polychotomous logistic regression was used to determine whether smoking was associated with an increased risk of being in a high-absence category.

The potential confounding variables (age, sex, race, drug use, job classification, and exercise) were simultaneously entered into the Cox and logistic regression models, permitting adjustment for confounding with cigarette smoking.[21] All *P* values were two sided and considered significant at a .05 level.

Results

Of the 4795 eligible applicants, 2537 were hired. There was no difference in smoking rate between those hired (32.5%) and not hired (33.2%; $P = .62$).

Table 1 presents the distribution of covariates. Smokers were more frequently older, female, nonexercisers, and drug users and less often Asian or letter carriers. Thus, these variables were potential confounders.

Table 2 shows measures related to the endpoints correlated with smoking status. Since the distribution for absence rate was skewed markedly to the right, both mean and median are provided. While this table does not include adjustments for covariates, inclusion of covariates in the Cox proportional hazards and polychotomous logistic regression models resulted in only small changes. Thus, these covariates were not major confounders in our data set.

Table 3 presents the results of the Cox proportional hazards models with inclusion of covariates. Cigarette smokers show increased risks of accidents, injuries, and discipline but no increased risk of termination.

A polychotomous logistic regression model was used to predict absence category with the covariates included. This analysis produced a beta for smokers of 0.264 with a *P* value of .0007. Interpretation of an odds ratio for absence is not

intuitive, as absence is divided into five leave-use categories. Mean absence rates provide a more easily interpreted estimate of the magnitude of the effect: 5.43% for smokers and 4.06% for nonsmokers.

Discussion

To our knowledge, this study presents the first prospective estimates of associations between cigarette smoking and adverse employment outcomes after adjustment for the potentially confounding covariates of age, sex, race, exercise, drug abuse, and job category. We found significant positive associations between cigarette smoking and absenteeism, industrial accidents, occupational injuries, and disciplinary action. Smokers had a 34% increase in mean absenteeism that remained significant after adjustment for covariates. Smokers had increased covariate-adjusted relative risks for industrial accidents (29%), occupational injuries (40%), and discipline (55%). We found no significant difference in turnover, voluntary or involuntary. Thus, our study confirms the previously reported elevated rates of absenteeism, accidents, and injuries for smokers.[3-9]

The mean age of our study population was 30.2 years. In our cohort of newly hired younger workers, any increase in absence from smoking-related illness is most likely due to acute, short-term illnesses such as respiratory infections. The more serious adverse health outcomes caused by cigarette smoking, cardiovascular and chronic pulmonary disease and cancers, probably would not develop until workers are several decades older. The increase in smoking-associated absenteeism in our cohort, therefore, probably underestimates the true magnitude of this effect for workers who smoke throughout their working lives. A similar analysis might be made for use of health insurance benefits, a factor that we did not assess but that has been reported to be increased in smokers.[6,7]

In this study, we were unable to control for the possible confounding effect of alcoholism. Smoking is reported to be associated with alcoholism.[13,14] Alcoholics are also known to have poor employment outcomes.[22] We did not conduct any laboratory tests for alcohol abuse, as the methodologies available lack adequate sensitivity and specificity. We also did not use self-reports, as surveys of alcohol use have consistently demonstrated underreporting, even in settings where there would be no adverse consequences for positive reports.[23] The validity of self-reported data would be highly questionable in the preemployment examination setting.

Since we did not control for the effects of alcohol abuse, it is possible that some of the associations demonstrated are attributable to alcoholism rather than smoking. The degree of any confounding by alcoholism would be a function of the coincidence of alcohol and tobacco abuse in our study population. It might, therefore, affect the degree to which the results of this research can be generalized to populations with different alcohol- and tobacco-abuse patterns.

Our study cannot distinguish adverse employment outcomes associated with cigarette smoking caused by the physiological effects of smoking from those caused by smoking behavior (e.g., decreased attention to driving) or by other social or personality characteristics associated with smoking. The postal service smoking policy in effect during the first 28 months of the study permitted either un-

TABLE 1—Demographic Characteristics of the Study Population

	Nonsmokers		Smokers		Total	
	No.	%	No.	%	No.	%
Sample size	1712	67.5	825	32.5	2537	100
Average time of follow-up (d)	407		404		406	
Mean age (y)	29.8		31.2		30.2	
Gender ($P < .001$)						
Male	1174	68.6	496	60.1	1670	65.8
Female	526	31.2	318	39.3	867	34.2
Race ($P = .001$)						
White	1521	88.8	757	91.8	2278	89.8
Black	107	6.3	53	6.4	160	6.3
Asian	84	4.9	15	1.8	99	3.9
Drug use ($P < .001$)						
None	1542	90.3	684	82.9	2226	87.9
Marijuana	103	6.0	94	11.4	197	7.8
Cocaine	33	1.9	22	2.7	55	2.2
Both	30	1.8	25	3.0	55	2.2
Exercise ($P < .001$)						
Aerobic	270	15.8	76	9.2	346	13.7
Nonaerobic	923	54.0	438	53.2	1361	53.8
None	516	30.2	310	37.6	824	32.6
Job classification ($P < .001$)						
Carrier	649	38.0	264	32.1	913	36.1
Clerk	121	7.1	44	5.4	165	6.5
LSM clerk	502	29.4	247	30.0	749	29.6
Mail handler	226	13.2	109	13.3	335	13.2
Mail processor	103	6.0	68	8.3	171	6.8
Other	109	6.4	90	10.9	199	7.9

Note. LSM = letter sorting machine. Since information on some covariates was not available for a few employees, sample sizes for some covariates may not add up to the overall totals. The "other" job classification includes special delivery messenger, motor vehicle operator, maintenance worker, computer forwarding system clerk, and electronic technician.

TABLE 2—Comparison of Endpoints between Smokers and Nonsmokers

	Nonsmokers, %	Smokers, %
Turnover (%)	19.6	21.5
Voluntary	12.8	13.0
Involuntary	6.4	8.4
Accidents (%)	18.5	22.6
Injuries (%)	10.9	15.3
Discipline (%)	7.9	12.1
Mean absence rate (%)	4.1	5.4
Median absence rate (%)	1.8	2.3

Note. Reported percentages represent the simple overall proportion of employees for whom the endpoint occurred; time- and covariate-adjusted analyses are reported in Table 3.

TABLE 3—Covariate-Adjusted Risks for Smokers Relative to Nonsmokers

	Relative Risk	*P* Value	95% Confidence Interval
Turnover	1.01	.956	0.83, 1.21
Voluntary	0.94	.585	0.74, 1.19
Involuntary	1.17	.306	0.86, 1.60
Accidents	1.29	.008	1.07, 1.55
Injuries	1.40	.004	1.11, 1.77
Discipline	1.55	.001	1.19, 2.02

Note. The risk estimates have been adjusted for age, gender, race, drug use, exercise habits, and job classification.

limited smoking at work or regular smoking breaks. For the last 3 months of follow-up, smoking was permitted during breaks every 2 hours. Thus, the effects of nicotine withdrawal would not have contributed to accidents or injuries. Although we cannot determine the etiology of the adverse employment outcomes, our findings provide a basis for estimating the magnitude of the increase in costs to industry that results from hiring cigarette smokers.

The findings of this study may also be useful in assessing the cost-effectiveness of corporate smoking control and smoking cessation programs. To the degree that such programs decrease cigarette smoking in the work force, a health-related effect such as absenteeism might be expected to decrease. It is less clear whether there would be similar improvements in effects such as accidents, injuries, and discipline, as these may be linked to social and behavioral characteristics that might persist after cessation of, or a decrease in, smoking. A comparison of smokers, former smokers, and persons who have never smoked might begin to answer such questions. We hope, however, that the present study will contribute to the body of knowledge needed to evaluate some of the complex ethical, legal, economic, and social issues related to smoking in the workplace. □

Acknowledgments

We are indebted to Dr. N. Sprince for critical review and comments; to C. Ryan for help with translation; to J. Paradise, T. Fenton, and E. Wright for help in data management; and to the staff of the US Postal Service Medical Unit in Boston for help in collecting the data. The opinions expressed in this paper, however, are those of the authors and do not necessarily reflect the views of the US Postal Service or those who so generously helped us.

References

1. US Department of Health, Education, and Welfare. *Smoking and Health. Report of the Advisory Committee to the Surgeon General of the Public Health Service.* Washington, DC: US Government Printing Office; 1964; PHS publication no. 1103.
2. US Department of Health and Human Services. *Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General.* Washington, DC: US Government Printing Office; 1989;267–284; DHHS publication no. (CDC) 89–8411.
3. Van Tuinen M, Land G. Smoking and excess sick leave in a department of health. *J Occup Med.* 1986;28:33–35.
4. Parkes KR. Smoking as a moderator of the relationship between affective state and absence from work. *J Appl Psychol.* 1983;68:698–708.
5. Hawker R, Holtby I. Smoking and absence from work in a population of student nurses. *Public Health (London).* 1988; 102:161–167.
6. Penner M, Penner S. Excess insured health care costs from tobacco-using employees in a large group plan. *J Occup Med.* 1990; 32:521–523.
7. Van Peenen PFD, Blanchard AG, Wolkonsky PM, Gill TM. Health insurance claims of petrochemical company employees. *J Occup Med.* 1986;28:237–240.
8. Naus A, Engler V, Hetychova M, Vavreckova O. Work injuries and smoking. *Ind Med Surg.* 1966;35:880–881.
9. Yuste PC, De Guevara ML. Influencia del fumar en los accidentes laborales. encuesta Estadistica. *Medicina y seguridad del trabajo.* 1973;21:38–46.
10. Cascio WF. Costing the effects of smoking in the work place. In: Cascio WF, *Costing Human Resources: The Financial Impact of Behavior in Organizations.* Boston, Mass: Kent Publishing Co; 1982;65–78.
11. Kristein MM. How much can business expect to profit from smoking cessation? *Prev Med.* 1983;12:358–381.
12. Whitehead PC, Smart RG, Laforest L. Multiple drug use among marijuana smokers in eastern Canada. *Int J Addict.* 1972;7:179–190.
13. Allen J, Mazzuchi J. Alcohol and drug abuse among American military personnel: prevalence and policy implications. *Milit Med.* 1985;150:250–255.
14. Dreher KF, Fraser JG. Smoking habits of alcoholic out-patients. *Int J Addict.* 1967;2:259–270.
15. Zwerling C, Ryan J, Orav EJ. The efficacy of preemployment drug screening for marijuana and cocaine in predicting employment outcome. *JAMA.* 1990;264:2639–2643.
16. Pierce JP, Dwyer T, DiGiusto E, et al. Cotinine validation of self-reported smoking in commercially run community surveys. *J Chronic Dis.* 1987;40:689–695.
17. Petitti DB, Friedman GD, Kahn W. Accuracy of information on smoking habits provided on self-administered research questionnaires. *Am J Public Health.* 1981;71: 308–311.
18. Bauman KE, Koch GG. Validity of self-reports and descriptive and analytical conclusions: the case of cigarette smoking by adolescents and their mothers. *Am J Epidemiol.* 1983;118:90–98.
19. Cox DR, Oakes D. *Analysis of Survival Data.* London: Chapman and Hall Ltd; 1984:91–109.
20. SAS Institute Inc. *SUGI Supplemental Library User's Guide.* Carey, NC: SAS Institute Inc; 1983.
21. Breslow NE, Day NE. *Statistical Methods in Cancer Research.* Lyons, France: International Agency for Research on Cancer; 1980;1:225–226.
22. Harwood HJ, Napolitano DM, Kristiansen PL, Collins JJ. *Economic Costs to Society of Alcohol and Drug Abuse and Mental Illness: 1980.* Research Triangle Park, NC: Research Triangle Institute; 1984. RTI/2734/00-01FR.
23. Pernanen K. Validity of survey data on alcohol use. In: Gibbins RJ, Israel Y, Kalant H, Popham RE, Schmidt W, Smart RG, eds. *Research Advances in Alcohol and Drug Problems.* New York, NY: John Wiley & Sons; 1974;1:355–374.

Misclassification of Smoking Status in the CARDIA Study: A Comparison of Self-report with Serum Cotinine Levels

Lynne E. Wagenknecht, DrPh, Gregory L. Burke, MD, Laura L. Perkins, PhD, Nancy J. Haley, PhD, and Gary D. Friedman, MD

ABSTRACT

Background. Although widely used in epidemiological studies, self-report has been shown to underestimate the prevalence of cigarette smoking in some populations.

Methods. In the CARDIA study, self-report of cigarette smoking was validated against a biochemical marker of nicotine uptake, serum cotinine.

Results. The prevalence of smoking was slightly lower when defined by self-report (30.9%) than when defined by cotinine levels equal to or greater than 14 ng/mL (32.2%, $P < .05$). The misclassification rate (proportion of reported nonsmokers with cotinine levels of at least 14 ng/mL) was 4.2% and was significantly higher among subjects who were Black, had a high school education or less, or were reported former smokers. Possible reasons for misclassification include reporting error, environmental tobacco smoke, and an inappropriate cutoff point for delineation of smoking status.

Using self-report as the gold standard, the cotinine cutoff points that maximized sensitivity and specificity were 14, 9, and 15 ng/mL for all, White, and Black subjects, respectively. The misclassification rate remained significantly higher in Black than in White subjects using these race-specific criteria.

Conclusions. Misclassification of cigarette smoking by self-report was low in these young adults; however, within certain race/education groups, self-report may underestimate smoking prevalence by up to 4%. (*Am J Public Health.* 1992;82: 33–36)

Introduction

Self-report is widely used to estimate the prevalence of cigarette smoking although it has been reported to underestimate the true prevalence in some populations.[1] Consequently, a number of biologic markers have been used to validate smoking, particularly in studies of cessation and of adolescent smoking.[2] Few studies have examined the validity of self-report in a general population sample of smokers and nonsmokers, nor in a sample of Blacks.

The purpose of this study is (1) to measure the discrepancy between smoking prevalence rates as defined by self-report and by serum cotinine, an objective measure of nicotine exposure, using a cotinine cutoff point of 14 ng/mL;[3,4] (2) to assess the degree of misclassification of reported nonsmokers by race, education, and past smoking behavior; and (3) to investigate possible reasons for misclassification.

Methods

The CARDIA study, a longitudinal epidemiological study of the risk factors for cardiovascular disease in a cohort of 5115 young adults aged 18 to 30 years, has been described previously.[5–7] The cohort was recruited to consist of approximately equal numbers of Black and White men and women of varied educational backgrounds. The data for this report were collected at the first clinical examination conducted in 1985 and 1986.

Smoking habit was assessed during the examination by interview. Smokers were defined as subjects who reported current, regular use of cigarettes (at least five cigarettes per week, almost every week). Nonsmokers were classified further as former smokers or as people who never smoked. Subjects who were reported being nonsmokers of cigarettes but who reported the current use of either cigars, pipe tobacco, smokeless tobacco, or nicotine gum were excluded from the analyses.

Smoking habit was also assessed at the recruitment telephone contact approximately 1 month prior to the examination. A smoker was defined as someone who smoked at least 100 cigarettes in his or her lifetime and who currently smoked. The reported number of hours per week that subjects were exposed to cigarette, cigar, or pipe smoke from others in their home was used as a measure of environmental tobacco smoke (ETS) exposure.

Serum cotinine was measured by radioimmunoassay[8,9] at the American Health Foundation in Valhalla, NY. Standard quality control procedures were followed; the interassay coefficient of variation was 7%. Based on previous investigations,[3,4] cotinine levels equal to or greater than 14 ng/mL were considered

Lynne E. Wagenknecht and Gregory L. Burke are in the Department of Public Health Sciences at the Bowman Gray School of Medicine in Winston-Salem, NC; Laura L. Perkins is with the School of Public Health at the University of Alabama at Birmingham; Nancy J. Haley is with the Metropolitan Life Insurance Company in Elmsford, NY; and Gary D. Friedman is with the Division of Research at Kaiser Permanente in Oakland, Calif.

Requests for reprints should be sent to Lynne E. Wagenknecht, DrPH, Department of Public Health Sciences, Bowman-Gray School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157-1063.

This paper was submitted to the journal December 26, 1990, and accepted with revisions May 28, 1991.

TABLE 1—Comparison of Smoking Prevalence Defined by Cotinine Versus Self-report, by Race and Education

Race/Education	n	Smoking Prevalence (%) Defined by Cotinine ≥ 14 ng/mL	Defined by Self-report	Difference	(95% CI)
Black ≤ HS	1317	47.1	43.8	3.3	(2.0,4.6)
Black > HS	1255	25.7	23.9	1.8	(0.8,2.9)
White ≤ HS	660	48.5	47.7	0.8	(–1.0,2.5)
White > HS	1752	19.6	19.8	–0.2	(–0.7,1.1)
Overall	4984[a]	32.2	30.9	1.3	(0.8,1.9)

Note. CI = confidence interval. ≤ HS = a high school education or less; > HS = more than a high school education.
[a]Thirty-five subjects were missing responses about cigarette smoking, 46 did not have cotinine results, and 48 reported use of other tobacco products, leaving 4984/5115 (97%) subjects available for analysis.

TABLE 2—Misclassification Rates, by Race and Education

Race/Education	All Reported Nonsmokers	Reported Never Smokers[c]	Reported Former Smokers
Black ≤ HS	8.2 (61/740)	5.1 (32/627)	25.7 (29/113)
Black > HS	3.7[aa] (35/955)	2.6[a] (22/841)	11.4[a] (13/114)
All Black	5.7 (96/1695)	3.7 (54/1468)	18.5 (42/227)
White ≤ HS	5.8 (20/345)	3.0 (7/230)	11.3[b] (13/115)
White > HS	2.1[aa,b] (29/1405)	1.0[a,b] (11/1092)	5.8[b] (18/313)
All White	2.8[bb] (49/1750)	1.4[bb] (18/1322)	7.2[bb] (31/428)
Overall	4.2 (145/3445)	2.6 (72/2790)	11.2 (73/655)

Note. Misclassification rate = proportion of reported nonsmokers with serum cotinine of at least 14 ng/mL. ≤ HS = a high school education or less; > HS = more than a high school education.
[a]$P < .05$, [aa]$P < .0001$ for difference in misclassification rates by education, within race.
[b]$P < .05$, [bb]$P < .0001$ for difference in misclassification rates by race, within education.
[c]All comparisons of misclassification rates by previous smoking history are significant, $P < .0001$.

TABLE 3—Distribution of Serum Cotinine Levels among Reported Nonsmokers and Smokers

Cotinine (ng/mL)	Reported Nonsmokers (n = 3445) n	Reported Nonsmokers (n = 3445) %[a]	Reported Smokers (n = 1539) n	Reported Smokers (n = 1539) %[a]
0–13	3300	—	78	—
14–49	72	49.7	133	9.1
50–99	23	15.9	142	9.7
100–149	15	10.3	206	14.1
150–199	7	4.8	197	13.5
200–249	8	5.5	220	15.1
250–299	9	6.2	151	10.3
300 +	11	7.6	412	28.2

[a]As a percent of all values of at least 14 ng/mL.

indicative of active smoking and not a result of ETS exposure.[10,11]

The Kolmogorov-Smirnov statistic was used to test for differences in the distributions of cotinine between Whites and Blacks and between former smokers and those who never smoked. Differences in ETS exposure between classes of nonsmokers were assessed using the Wilcoxon rank sum test. Confidence intervals were computed for differences between smoking prevalence rates using variance formulas for rate differences and taking the matched pairs into account.[12]

Results

Overall, the smoking prevalence rate as defined by cotinine levels of at least 14 ng/mL was 1.3 percentage points higher than that based on self-report (Table 1). The largest discrepancies were observed among the Black subjects; self-report underestimated smoking prevalence by 3.3 percentage points among Blacks with a high school education or less.

Of the 3445 reported nonsmokers, 145 had cotinine levels of at least 14 ng/mL, for a misclassification rate of 4.2% (Table 2). (The proportion of reported smokers with cotinine levels of less than 14 ng/mL was 78/1539, or 5.1%.) The misclassification rate was two times greater among Black subjects than among White subjects (5.7% vs 2.8%), almost three times greater among those with a high school education or less than among those with more than a high school education (7.5% vs 2.8%), and four times higher among the reported former smokers than among those who reported never having smoked (11.2% vs 2.6%, all $P < .0001$). Gender differences (not shown) were negligible.

The distribution of cotinine levels among the 145 subjects misclassified as nonsmokers differed from that among the true smokers, i.e., reported smokers with cotinine levels of at least 14 ng/mL (Table 3). One half of the misclassified subjects had cotinine levels between 14 and 49 ng/mL whereas less than 10% of the true smokers had cotinine levels in this range. The distribution of cotinine levels equal to or greater than 14 ng/mL was not significantly different either between former smokers and those who never smoked or between the races (not shown).

One half (73/144) of the subjects misclassified as nonsmokers reported spending most of their time around smokers whereas only 28% (906/3296) of the true nonsmokers (reported nonsmokers with cotinine of less than 14 ng/mL) reported this behavior ($P < .0001$). This was confirmed by their self-report of the number of hours they were exposed to ETS. The misclassified nonsmokers reported greater weekly exposure to ETS than the true nonsmokers (medians: 5 hours vs 0 hours, $P < .0001$); among the misclassified subjects, Blacks reported greater weekly exposure to ETS than Whites (medians: 6 hours vs 2 hours, $P = .02$). In addition, nearly one half (29/67) of the misclassified former smokers reported having quit smoking within the past year whereas only 20% of the true reported former

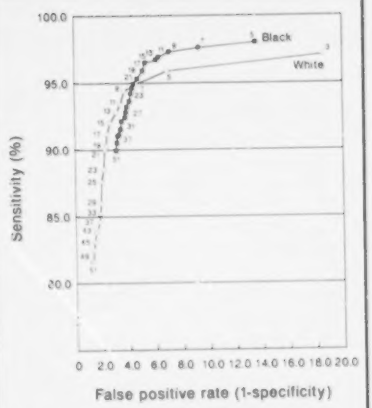


FIGURE 1—Receiver operating characteristic curve for varying levels of serum cotinine by race. The false positive rate (i.e., the proportion of reported nonsmokers whose cotinine level is equal to or greater than the given cutoff point) is the same as the misclassification rate as defined in this article. Levels of cotinine equal to or above the cutoff point are considered characteristic of a smoker. Self-report was considered truth for the calculation of these values.

TABLE 4—Sensitivity (%) and False Positive Rate (1-Specificity) for Various Cotinine Cutoff Values (ng/mL)

Cotinine Cutoff	Overall			Black			White		
	Sens.	False Pos. (1-Spec.)	Sum[a]	Sens.	False Pos. (1-Spec.)	Sum[a]	Sens.	False Pos. (1-Spec.)	Sum[a]
5	97.1	10.2	186.9	98.0	13.6	184.4	95.9	6.9	189.0
7	96.4	6.9	189.5	97.6	9.3	188.3	94.9	4.6	190.3
9	96.0	5.4	190.6	97.3	7.1	190.2	94.4	3.7	190.7[b]
11	95.4	4.9	190.5	96.9	6.3	190.6	93.4	3.4	190.0
13	95.0	4.6	190.4	96.7	6.1	190.6	92.8	3.1	189.7
14	94.9	4.2	190.7[b]	96.6	5.7	190.9	92.8	2.8	190.0
15	94.5	4.0	190.5	96.5	5.3	191.2[b]	92.0	2.6	189.4
17	93.9	3.7	190.2	95.9	5.1	190.8	91.2	2.4	188.8
19	93.2	3.5	189.7	95.3	4.7	190.6	90.3	2.3	188.0

Note. Cotinine levels < cutoff indicate nonsmoker, levels ≥ cutoff indicate smoker. Number of reported nonsmokers: 1695 Blacks, 1750 Whites; number of reported smokers: 877 Blacks, 662 Whites.
[a]Sum = sensitivity + specificity.
[b]Point at which sensitivity and specificity are maximized. Note that only odd values of cotinine are presented except for the value of 14 ng/mL.

smokers reported having quit over that period ($P < .0001$). Note that denominators are reduced slightly due to missing values.

Approximately one third (42/145) of the subjects misclassified as nonsmokers had identified themselves as current smokers at the recruitment interview one month earlier. Misclassification rates, calculated using smoking status as obtained at recruitment, were slightly higher than and yet consistent with those presented in Table 1; rates were twice as high in Black subjects than in White subjects (8.8% vs 3.8%) and three times higher in the reported former smokers than in those who reported having never smoked (14.5% vs 4.5%).

To investigate the possibility that the misclassification of these 145 reported nonsmokers was due to the use of an inappropriate cotinine cutoff point, we examined the receiver operating characteristic curve, which plots sensitivity against the false positive rate[13] (Figure 1). Cutoff points in the upper left corner of the graph maximize sensitivity and specificity. For these analyses, self-report was considered truth.

The cotinine cutoff point that led to the best combined levels of sensitivity and specificity in the overall cohort was 14 ng/mL (Table 4; note that sensitivity and specificity are both quite high over a range of cotinine cutoff points, from 9 ng/mL to 16 ng/mL). The race-specific cutoff point that maximized these parameters in the Black subjects was slightly higher than that in the White subjects (Figure 1, Table 4). (This was true across the entire range of sensitivity and specificity.) Among Black subjects, a cutoff point of 15 ng/mL maximized these parameters, yielding a sensitivity of 96.5% and a specificity of 94.7% (false positive rate of 5.3%). (Note that the definition of the false positive rate is the same as that of the misclassification rate). Among White subjects, a cutoff point of 9 ng/mL maximized these parameters, yielding a sensitivity of 94.4% and a specificity of 96.3% (false positive rate of 3.7%).

Applying these race-specific cutoff points to the entire cohort, the overall misclassification rate for reported nonsmokers was 4.5% (155/3445); these rates remained higher in Black subjects, as shown above (5.3% vs 3.7%, $P < .05$). Using these race-specific cutoff points to define smoking prevalence led to findings similar to those reported in Table 1, with a significant yet smaller underestimation of prevalence among Blacks (and also a significant underestimation among Whites with a high school education or less).

Discussion

These data indicate that self-report is a fairly accurate estimator of smoking prevalence in young adults. However, judging by cotinine levels and using a cutoff point of 14 ng/mL, smoking prevalence is underestimated to a greater degree among Blacks than among Whites, among former smokers than among those who never smoked, and among those with a high school education or less than among those with more.

Four possible reasons for misclassification were considered: (1) data collection/entry errors, (2) reporting errors, (3) ETS exposure, and (4) use of an inappropriate cutoff point for delineation of smoking status. In response to the first possibility, data collection followed standard quality control procedures. No data entry errors were found.

As for the second possibility, several types of reporting errors could have occurred. First, the subjects may have misunderstood the smoking questions. In fact, 29% of the 145 misclassified subjects identified themselves as smokers at recruitment. (It is possible that some of these subjects quit smoking in the month between the initial telephone contact and their CARDIA examination; however, we were unable to assess this.) Perhaps the questions asked at recruitment more readily identified the "light" smokers. Indeed, a large number of the misclassified subjects were light smokers, based on their cotinine levels. Considering this explanation, it was surprising that the misclassification rate was higher when comparing cotinine levels with smoking status reported at recruitment.

Differential reporting error could explain the higher misclassification rate among Black subjects. Compared with

Whites, Blacks may underreport smoking to a greater degree because they are, in general, lighter smokers.[6] However, when the smoking status reported at recruitment (the more liberal definition) was compared with serum cotinine levels, the difference in misclassification rates between White and Black subjects persisted.

Another explanation of misclassification, also a reporting error, is denial of smoking. Study subjects may wish to report behaviors consistent with a healthy life-style. Responses received in the possibly less-threatening recruitment telephone interview support this explanation. Denial may be an important factor particularly among the former smokers, many of whom reported having quit within the past year. A desire to "kick the habit" may lead these reported former smokers to underreport their habit to the apparent marked degree noted.

Furthermore, misclassification may partially be the result of an inappropriate cotinine cutoff point. However, this does not appear to be the case; 14 ng/mL was identified as the "best" cutoff point for the overall sample. This concern is more relevant to assessing the possible cause of higher misclassification rates observed among Black subjects. Race-specific cotinine cutoff points seemed necessary because of our earlier research, which showed significantly greater cotinine levels among Black smokers than among White ones over all levels of reported nicotine exposure.[7] Furthermore, the study that originally proposed the use of a cotinine cutoff point of 14 ng/mL was conducted with a small sample of White subjects and may not be generalizable to other samples.[4] However, the race-specific cutoff points that optimized sensitivity and specificity were not considerably different from 14 ng/mL, nor did they correct the higher misclassification rates in Blacks.

Our use of self-report as the gold standard for determining the best cotinine cutoff point seems incompatible with our initial challenge of the use of questionnaire data to assess smoking prevalence. Nevertheless, a measure of internal validity suggests that self-report is reasonably valid (i.e., the smoking status of 95% of the subjects was concordant in the tobacco interview and recruitment even though some discordance was expected).

Although, in general, exposure to ETS is an unlikely explanation for misclassification, it should not be dismissed in all cases. Reported ETS exposure was significantly higher among the misclassified subjects and among Blacks. This may partially explain higher rates of misclassification among Black subjects. Differences in nicotine metabolism (i.e., longer elimination times for nicotine and its metabolites even among nonsmokers) should also be considered in further investigations of this racial difference in cotinine levels.[7,14]

In summary, these data suggest that self-report is a reasonably accurate indicator of smoking prevalence in a biracial cohort of young adults. Misclassification is more pronounced among Black subjects, former smokers, and those with a high school education or less. Higher misclassification rates among Black subjects could not be explained by the choice of a cotinine cutoff point; however, reporting errors may account for higher misclassification rates among the former smokers and the less educated. Although we identified slightly different race-specific cutoff points than the previously proposed cutoff point of 14 ng/mL, these points had little effect on our findings. □

Acknowledgments

This study was supported by National Heart, Lung, and Blood Institute contracts N01-HC-48047, N01-HC-48048, N01-HC-48049, and N01-HC-48050. This paper was presented in part at the 30th Annual Conference on Cardiovascular Disease Epidemiology in San Diego in March 1990.

References

1. Coultas DB, Howard CA, Peake GT, Skipper BJ, Samet JM. Discrepancies between self-reported and validated cigarette smoking in a community survey of New Mexico Hispanics. *Am Rev Respir Dis.* 1988;137:810–814.
2. Stookey GK, Katz BP, Olson BL, Drook CA, Cohen SJ. Evaluation of biochemical validation measures in determination of smoking status. *J Dent Res.* 1987;66:1597–1601.
3. Cummings SR, Richard RJ. Optimum cutoff points for biochemical validation of smoking status. *Am J Public Health.* 1988;78:574–575.
4. Jarvis MJ, Tunstall-Pedoe H, Feyerabend C, Vesey C, Saloojee Y. Comparison of tests used to distinguish smokers from nonsmokers. *Am J Public Health.* 1987;77:1435–1438.
5. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol.* 1988;41:1105–1116.
6. Wagenknecht LE, Perkins LL, Cutter GR, et al. The impact of education on smoking in young adults: the CARDIA study. *Prev Med.* 1990;19:158–169.
7. Wagenknecht LE, Cutter GR, Haley NJ, et al. Racial differences in serum cotinine levels among smokers in the CARDIA study. *Am J Public Health.* 1990;80:1053–1056.
8. Haley NJ, Axelrad CM, Tilton KA. Validation of self-reported smoking behavior: biochemical analyses of cotinine and thiocyanate. *Am J Public Health.* 1983;73:1204–1207.
9. VanVunakis H, Gjika HB, Langone JJ. Method 16–radioimmunoassay for nicotine and cotinine. In: O'Neill IK, Brunnemann KD, Dodet B, Hoffmann D, eds. *Environmental Carcinogens Methods of Analysis and Exposure Measurement.* Vol. 9 *Passive Smoking.* 1987;317–330. IARC Scientific Publication, No. 81. Oxford University Press, NY.
10. Benowitz NL. The use of biologic fluid samples in assessing tobacco smoke consumption. In: Grabowshi J, Bell C, eds. *Measurement in the Analysis of Treatment of Smoking Behavior.* Rockville, Md: National Institute on Drug Abuse; 1983:6–26. NIDA Research Monograph No. 48.
11. Hoffmann D, Haley NJ, Adams JD, Brunnemann KD. Tobacco sidestream smoke: uptake by nonsmokers. *Prev Med.* 1984;13:608–617.
12. Rothman KJ. *Modern Epidemiology.* Boston, Mass: Little, Brown and Co; 1986:276.
13. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med.* 1978;8:283–298.
14. Sepkovic DW, Haley NJ, Hoffmann D. Elimination from the body of tobacco products by smokers and passive smokers. *JAMA.* 1986;256:863. Letter.

A Prospective Study of the Attributable Risk of Cancer Due to Cigarette Smoking

Po-Huang Chyou, PhD, Abraham M. Y. Nomura, MD, and Grant N. Stemmermann, MD

ABSTRACT

Background. The goals of this study were to measure the impact of cigarette smoking on cancer incidence and to determine the attributable risk of cancer due to smoking.

Methods. A cigarette smoking history was obtained from 8006 Japanese-American men examined from 1965 through 1968. After 22 years, 1389 incident cases of cancer were identified. There were 212 men with lung cancer; 202 with oral, esophageal, laryngeal, pancreatic, renal, ureteral, or bladder (oral–bladder) cancer; and 975 with cancer at other sites.

Results. Current smokers at time of examination had a higher incidence than nonsmokers for each of the three cancer site categories. Eighty-five percent of lung cancer cases diagnosed among current and never smokers can be attributed to cigarette smoking. The attributable risks were 46%, 16%, and 29%, respectively, for oral–bladder cancers, other cancers, and all cancers combined. In turn, the corresponding attributable risks were 60%, 26%, 13%, and 21% in comparing current smokers with past smokers.

Conclusions. Current smokers can greatly reduce their risk of cancer, especially lung cancer, if they quit smoking. (*Am J Public Health.* 1992;82:37–40)

Introduction

Cancer is the second leading cause of death in the United States. It was estimated, for 1991, that more than 1 million new cancer cases and 514 000 cancer deaths would occur in the country.[1] The surgeon general of the US Public Health Service has reported that cigarette smoking is a major cause of cancer mortality in the US population.[2]

Eight large prospective studies in five countries have provided much of the information on the association between tobacco and cancer.[3–10] These mortality studies have consistently reported that cigarette smoking increases the risk of cancer deaths, especially lung cancer. They also have found that cigarette smoking leads to oral, laryngeal, and esophageal cancer and is associated with cancer of the pancreas, urinary bladder, and kidney.[2]

None of these prospective studies reported results on the incidence of cancer and the attributable risk associated with cigarette smoking. Because some subjects afflicted with cancer do not succumb to their disease, it is also important to study the impact of smoking on cancer incidence. We had the opportunity to determine the association of cancer incidence with cigarette smoking in a large cohort of 8006 men and to compute the attributable risk due to cigarette smoking. The attributable risk provides an estimate of the extent to which the excess cancer risk might be reduced if cigarette smoking were discontinued or not initiated in a study population.

Methods

American men of Japanese ancestry, born in the years 1900 through 1919 and living on the Hawaiian island of Oahu, were identified by the Honolulu Heart Program in 1965.[11] Of the 11 148 identified men, 8006 (71.8%) were interviewed and examined between 1965 and 1968, 180 (1.6%) died before they could be examined, and 2962 (26.6%) did not participate in the program. Of the 8006 men examined, the following were excluded from this analysis: 81 men with prevalent cases of cancer, 164 men with a diagnosis of cancer not confirmed by histology, and 1 subject in whom the smoking status was unknown. A total of 7760 men remained.

During the interview, information on cigarette smoking was obtained. Current smokers recorded the usual number of cigarettes smoked per day and the number of years they had smoked. Past smokers recorded the age they started smoking, the maximum number of cigarettes smoked per day, the number of years smoked, and the age they stopped smoking.

Since examination, newly diagnosed cases of cancer have been identified through continuous surveillance of all general hospitals on Oahu. From time of examination to May 1990, the study identified 1389 incident cancer cases (*International Classification of Diseases*, 8th version, codes 140–209) with tissue confirmation of their diagnosis. There were 212 men with lung cancer (code 162.1); 104 with renal, ureteral, or bladder cancer (codes 189 and 188); 35 with pan-

The authors are with the Japan-Hawaii Cancer Study, Kuakini Medical Center, Honolulu.

Requests for reprints should be sent to Po-Huang Chyou, PhD, Japan-Hawaii Cancer Study, Kuakini Medical Center, 347 North Kuakini Street, Honolulu, HI 96817.

This paper was submitted to the journal January 29, 1991, and accepted with revisions July 10, 1991.

TABLE 1—Percentage of Smokers and Nonsmokers by Age Groups

Age at Exam	Smoking Status at Time of Interview: Never No.	Never %	Past No.	Past %	Current No.	Current %	Total
45–49	476	26.1	475	26.0	876	47.9	1827
50–54	825	30.3	667	24.5	1229	45.2	2721
55–59	497	32.3	423	27.5	617	40.2	1537
60–64	423	33.4	341	26.9	504	39.7	1268
65+	123	30.2	125	30.7	159	39.1	407
Total	2344	30.2	2031	26.2	3385	43.6	7760

creatic cancer (code 157); 25 with esophageal cancer (code 150); 24 with laryngeal cancer (code 161); 14 with oral-cavity cancer (codes 140–146); and 975 with cancer at other sites.

On the basis of a separate 19-year follow-up survey of the study subjects since their examination, we found that only 1.3% of the men could not be located on Oahu. However, these men were kept in the study because some of them could still be hospitalized on Oahu later even if they had left the island, and others could still be on Oahu even though their whereabouts were unknown.

The lung cancer cases were analyzed separately, while other cancers that have been related to cigarette smoking were combined because of their smaller numbers. These included cancer of the oral cavity, larynx, esophagus, pancreas, kidney, ureter, and bladder. They were designated as oral–bladder cancers for ease of presentation.

Site-specific and total cancer incidence rates for never, past, and current smokers were adjusted for age at examination via the direct method of standardization, with the entire cohort serving as the reference population. The reference population is the same as the one used to provide age standardization. The relative risks (RR) and 95% confidence intervals for the above cancers associated with cigarette smoking were derived from Cox's proportional hazards regression models[12] while adjusting for age at exam. Separate likelihood ratio tests were performed to evaluate the linear trend associations between cancer risks for each site and current smoking in pack-years (treated as a continuous variable).

To determine the percentage of cancer risk attributable to smoking among current and never smokers, we calculated the population attributable risk (AR) using Bruzzi's method[13] with adjustment for age. The estimated AR is similar to the incidence–density fraction of Greenland and Robins.[14] Similar analysis was done for current smokers compared with past smokers.

Results

Table 1 shows the percentage of current, past, and never smokers by 5-year age groups. The percentage of current smokers decreased with age.

The age-adjusted incidence rates and relative risks for specific cancers and all cancers by smoking status are presented in Table 2. The rates for past smokers were consistently higher than those of nonsmokers. However, the highest rates were found among current smokers. The increase in relative risk among current smokers was statistically significant ($P < .05$) for each comparison with nonsmokers.

Table 3 gives the age-adjusted relative risk of cancer by site for current smokers according to their pack-years of cigarette smoking. There was a significantly positive trend for each site by increasing amount of cigarette use. The relative risk of lung cancer was 23.3 for those who had smoked more than 45 pack-years.

The age-adjusted attributable risks of cancer for current smokers at time of interview compared with nonsmokers and past smokers are presented in Table 4. With adjustment for age, 85% of the total number of lung cancer cases among current and never smokers can be attributed to cigarette smoking. The attributable risks were 46%, 16%, and 29%, respectively, for oral–bladder cancers, other cancers, and all cancers combined. In turn, 60% of the total number of lung cancers among current and past smokers can be attributed to the persistence in smoking. The attributable risks were, correspondingly, 26% for oral–bladder cancers, 13% for other cancers, and 21% for all cancers.

Discussion

All Cancers

The incidence data in this study showed that current and past smokers at time of interview had a significantly greater risk for all cancers than their nonsmoking counterparts. The observed relative risks were 1.9 for current smokers and 1.2 for past smokers. These results were similar to those of cancer mortality studies that found that the overall cancer mortality risks ranged from 1.7 to 2.6 for current cigarette smokers and from 1.4 to 1.5 for past smokers.[5,15,16]

A principal observation in this study is that 29% of total cancer risk among current and never smokers can be attributed to cigarette smoking. It was also observed that total cancer risk among smokers could be reduced by 21% if current smokers became past smokers, as shown in Table 4. These data are encouraging and suggest that the cessation of cigarette smoking can reduce the risk of cancer. The 1982 surgeon general's report noted that smoking contributed to 43% of all US male cancer deaths.[2] However, the report neither separated the effect among current and past smokers nor adjusted for age. If current and past smokers are combined in our study, 26% of all incident cancer cases can be attributed to cigarette smoking.

Lung Cancer

Lung cancer is the site with the highest relative risk due to cigarette smoking. Current smokers at time of interview had a relative risk of 12, while a threefold increase in risk for lung cancer was also observed for past smokers. In addition, a strong dose–response trend was observed with increasing number of pack-years among current smokers.

Our findings are similar to those of the 16-year follow-up study of US veterans.[5] It reported that the mortality ratios for lung cancer were 11.3 for current smokers and 4.0 for past smokers. Seven other prospective studies[3,4,6–10] and three case–control studies[17–19] showed a consistently positive pattern between cigarette smoking and lung cancer, with a relative risk up to 14.2 in the Canadian Veterans Study.[6]

Our study observed that 85% of lung cancer incidence among current and never smokers could have been avoided if current smokers had never smoked cigarettes. If current smokers had ceased smoking, 60% of the lung cancer risk that occurred within the population of current

TABLE 2—Age-adjusted Incidence Rates per 1000 Men, and Relative Risks (RR) for Specific Cancer Sites, by Smoking Status at Time of Interview

		Smoking Status									
		Never			Past				Current			
Cancer Site	No.	No.	Rate	RR	No.	Rate	RR	95% CI	No.	Rate	RR	95% CI
Lung	212	12	5.9	1.0	31	18.0	3.2	1.6, 6.2	169	62.5	12.0	6.7, 21.6
Oral–Bladder[a]	202	37	17.6	1.0	49	27.8	1.6	1.1, 2.5	116	42.9	2.7	1.8, 3.9
Other	975	270	116.1	1.0	243	124.0	1.1	0.9, 1.3	462	154.1	1.5	1.3, 1.7
All	1389	319	133.8	1.0	323	157.8	1.2	1.0, 1.4	747	225.6	1.9	1.7, 2.2
Noncases	6371	2025			1708				2638			

Note. CI = confidence interval.
[a]Cancer of the oral cavity, larynx, esophagus, pancreas, kidney, ureter, or bladder.

TABLE 3—Age-adjusted Relative Risks (RR) for Specific Cancers and All Cancers, by Current Smoking in Pack-Years at Time of Interview

			Current Smoking (pack-years)									
	Never smoked		<31			31–45			46+			P-value for trend
Cancer Site	No.	RR	No.	RR	95% CI	No.	RR	95% CI	No.	RR	95% CI	
Lung	12	1.0	33	6.3	3.3, 12.3	44	9.0	4.8, 17.1	92	23.3	12.8, 42.6	<.0001
Oral–Bladder[a]	37	1.0	37	2.3	1.4, 3.6	42	2.8	1.8, 4.3	36	3.0	1.9, 4.8	<.0001
Other	270	1.0	138	1.2	1.0, 1.5	182	1.6	1.4, 2.0	136	1.6	1.3, 2.0	<.0001
All	319	1.0	208	1.5	1.3, 1.8	268	2.0	1.7, 2.3	264	2.4	2.1, 2.9	<.0001
Noncases	2025		996			880			738			

Note. Current smokers whose smoking history in pack-years were unknown were excluded. CI = confidence interval.
[a]Cancer of the oral cavity, larynx, esophagus, pancreas, kidney, ureter, or bladder.

and past smokers could have been avoided. For smokers in a case–control study in Louisiana, 90% of the lung cancer risk was attributable to active cigarette smoking.[19] The 1982 surgeon general's report estimated that 91% of lung cancer deaths were attributable to the effects of smoking.[2]

Oral–Bladder Cancers

The risk of oral–bladder cancers among current smokers at time of interview was significantly elevated in this study. An accentuation in risk was observed with increasing number of pack-years of cigarette smoking among current smokers.

Wynder et al. have noted that smoking-related cancers in men include cancer of the mouth, larynx, esophagus, pancreas, and bladder, in addition to the lung.[20] In four earlier studies,[3,15,21,22] the mortality risks for the combination of oral, pharyngeal, laryngeal, and esophageal cancer ranged from 2.2 to 7.0 for smokers compared with nonsmokers. Other studies have reported results separately for cancer of the pancreas,[5,22–26] bladder,[5,22,26–28] and kidney.[5,22,29] The risks for these cancers ranged from 1.4 to 5.4 for current smokers, 1.2 to 1.6 for past smokers, and 1.4 to 2.7 for smokers as a group, compared with nonsmokers.

TABLE 4—Age-adjusted Attributable Risks for Specific Cancers and All Cancers for Current Smokers at Time of Interview Compared with Never Smokers and Past Smokers

Cancer Site	Current Smokers vs Never Smokers (%)	Current Smokers vs Past Smokers (%)
Lung	85	60
Oral–bladder[a]	46	26
Other	16	13
All	29	21

[a]Cancer of the oral cavity, larynx, esophagus, pancreas, kidney, ureter, or bladder.

The 1982 surgeon general's report noted that the percentages of death among US males possibly due to smoking were 75% for mouth, pharyngeal, laryngeal, or esophageal cancer; 56% for bladder cancer; and 40% for pancreatic cancer. Our findings showed that 46% of the incidence due to oral–bladder cancers among current and never smokers could be avoided if current smokers had never smoked. Alternatively, if they give up smoking, their excess risk for these cancers could be reduced by 26%.

Other Cancers

Current smokers had a modest risk (RR = 1.5) for other cancers in the study. A positive trend in risk with increasing pack-years of smoking among current smokers was also noted, but the dose–response trend was not linear. If there is a true increase in risk of other cancers due to cigarette smoking, it is not as substan-

tial as that for lung and oral–bladder cancers. The data suggest that up to 16% of the incidence of other cancers among current and never smokers could be avoided if current smokers had never smoked.

Limitations of the Study

The assumption was made in this study that a person's smoking status at time of interview would remain the same throughout the follow-up period. Other studies,[3,5] however, have indicated that some men who reported themselves as current cigarette smokers at the beginning of the study had stopped smoking during the course of the study. Consequently, the group of current smokers included an appreciable proportion of ex-cigarette smokers. If such misclassifications/biases existed in our study, the reported risks of cancer for current smokers would probably be underestimated, suggesting that the actual benefits of smoking cessation could be greater than those shown. It is very unlikely that nonsmoking subjects, who were 45 through 68 years of age at time of interview, started smoking during the course of the study.

Conclusion

The present study provides cancer incidence data consistent with past cancer mortality data of the harmful effects of cigarette smoking. Evidence is especially overwhelming on the association between current cigarette smoking and the risk of lung cancer. An effective way of reducing the risk of cancer associated with cigarette smoking is to quit smoking, as shown by the data on attributable risk. Development of smoking-cessation programs that would enable current smokers to permanently quit their cigarette habit within the shortest time period should be strongly supported. □

Acknowledgments

This research was supported by Grant R01 CA 33644 from the National Cancer Institute, National Institutes of Health.

References

1. Boring CC, Squires TS, Tong T. Cancer statistics, 1991. CA. 1991;41:19–36.
2. US Department of Health and Human Services. *The Health Consequences of Smoking. Cancer: A Report of the Surgeon General.* Rockville, Md: US Department of Health and Human Services; 1982.
3. Doll R, Peto R. Mortality in relation to smoking: 20 years' observations on male British doctors. *Br Med J.* 1976;2:1525–1536.
4. Hammond EC, Horn D. Smoking and death rates—report on forty-four months of follow-up of 187,783 men. II. Death rates by cause. *JAMA.* 1958;166:1294–1308.
5. Rogot E, Murray JL. Smoking and causes of death among U.S. veterans: 16 years of observation. *Public Health Rep.* 1980;95:213–222.
6. Best EWR, Josie GH, Walker CB. A Canadian study of mortality in relation to smoking habits: a preliminary report. *Canadian J Public Health.* 1961;52:99–106.
7. Weir JM, Dunn JE Jr. Smoking and mortality: a prospective study. *Cancer.* 1970;25:105–112.
8. Hammond EC. Smoking in relation to the death rates of one million men and women. In: Haenszel W, ed. *Epidemiological Approaches to the Study of Cancer and Other Chronic Diseases.* Natl Cancer Inst Monogr No. 19. Rockville, Md: US Department of Health, Education, and Welfare; 1966:127–204.
9. Hirayama T. Smoking in relation to the death rates of 265,118 men and women in Japan. A report on 5 years of follow-up. Presented at the American Cancer Society's 14th Science Writers' Seminar; March 24–29; Clearwater Beach, Fla.
10. Cederlof R, Friberg L, Hrubec Z, Lorich U. *The Relationship of Smoking and Some Social Covariables to Mortality and Cancer Morbidity. A Ten Year Follow-up in a Probability Sample of 55,000 Swedish Subjects Age 18 to 69. Part I and II.* Stockholm, Sweden: Karolinska Institute; Department of Environmental Hygiene; 1975.
11. Worth R, Kagan A. Ascertainment of men of Japanese ancestry in Hawaii through World War II Selective Service registration. *J Chronic Dis.* 1970;23:389–397.
12. Cox DR. Regression models and life tables (with discussion). *J Royal Stat Soc, Series B.* 1972;34:187–220.
13. Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C. Estimating the population attributable risk for multiple risk factors using case-control data. *Am J Epidemiol.* 1985;122:904–914.
14. Greenland S, Robins JM. Conceptual problems in the definition and interpretation of attributable fractions. *Am J Epidemiol.* 1988;128:1185–1197.
15. Dorn HF. Tobacco consumption and mortality from cancer and other diseases. *Public Health Rep.* 1959;74:581–593.
16. Friedman GD, Dales LG, Ury HK. Mortality in middle-aged smokers and nonsmokers. *N Engl J Med.* 1979;300:213–217.
17. Damber LA, Larsson L-G. Smoking and lung cancer with special regard to type of smoking and type of cancer. A case-control study in north Sweden. *Br J Cancer.* 1986;53:673–681.
18. Stayner LT, Wegman DH. Smoking, occupation, and histopathology of lung cancer: a case-control study with the use of the Third National Cancer Survey. *JNCI.* 1983;70:421–426.
19. Fontham ETH, Correa P, Chen VW, Craig JF, Pickle LW, Falk R. Tobacco and cancer. *J Louisiana State Med Soc.* 1988;140(4):29–30, 35–40.
20. Wynder EL, Hoffman D. Tobacco. In: Schottenfeld D, Fraumeni JF, eds. *Cancer Epidemiology and Prevention.* Philadelphia: WB Saunders Company; 1982:277–292.
21. Hammond EC, Seidman H. Smoking and cancer in the United States. *Prev Med.* 1980;9:169–173.
22. Doll R, Peto R. The causes of cancer. Oxford Medical Publications. *JNCI.* 1981;66:1191–1308.
23. Hirayama T. Epidemiology of pancreatic cancer in Japan. *Int J Pancreatology.* 1988;3:S203–S204.
24. Mills PK, Beeson WL, Abbey DE, Fraser GE, Phillips RL. Dietary habits and past medical history as related to fatal pancreas cancer risk among Adventists. *Cancer.* 1988;61:2578–2585.
25. MacMahon B, Yen S, Trichopoulos D, Warren K, Nardi G. Coffee and cancer of the pancreas. *N Engl J Med.* 1981;304:630–633.
26. Whittemore AS, Paffenbarger RS Jr, Anderson K, Lee JE. Early precursors of site-specific cancers in college men and women. *JNCI.* 1985;74:43–51.
27. Cole P, Monson RR, Haning H, Friedell GH. Smoking and cancer of the lower urinary tract. *N Engl J Med.* 1971;284:129–134.
28. Brownson RC, Chang JC, Davis JR. Occupation, smoking, and alcohol in the epidemiology of bladder cancer. *Am J Public Health.* 1987;77:1298–1300.
29. Wynder EL, Mabuchi K, Whitmore WF Jr. Epidemiology of adenocarcinoma of the kidney. *JNCI.* 1974;53:1619–1634.

Brief Supportive Telephone Outreach as a Recruitment and Intervention Strategy for Smoking Cessation

Harry A. Lando, PhD, Wendy L. Hellerstedt, MPH, Phyllis L. Pirie, PhD, and Paul G. McGovern, PhD

ABSTRACT

Background. Formal efforts to recruit smokers into cessation programs have failed to reach large segments of the smoking population. Telephone intervention may represent a viable strategy to promote smoking cessation. An even more promising approach may be a combination of brief telephone support and outreach to identified smokers.

Methods. Telephone support for smoking cessation was provided to four identified smoker populations in Bloomington, Minn, one of three Minnesota Heart Health Program education communities. Smokers were randomly assigned to an intervention consisting of two 15-minute telephone calls approximately 1 to 3 weeks apart or to a nonintervention control.

Results. At the 6-month follow-up, a significant overall effect was found in favor of the intervention condition for both self-reported and cotinine-validated quitting. Differences between intervention and control conditions were no longer significant at 18 months.

Conclusions. Smokers' receptivity to telephone intervention was at least moderately encouraging. The cost of intervention could be relatively low if trained volunteers initiated telephone calls. However, more intensive telephone intervention and support may be needed to produce lasting changes in smoking prevalence. (*Am J Public Health.* 1992; 82:41–46)

Introduction

Formal efforts to recruit smokers into cessation programs have failed to reach large segments of the smoking population.[1,2] Although the vast majority of smokers express a desire to quit, they also state a preference to quit on their own.[1,3] Indeed, the vast majority of smokers who quit do so without formal assistance.[4] Unfortunately, individual unaided quit attempts tend to have a low long-term success rate; reports indicate that fewer than 10% of those who quit on their own maintain abstinence for one year.[5,6]

Telephone intervention may represent a viable strategy to promote smoking cessation. Such intervention is distinct from telephone help lines, which function less as proactive interventions than as resources for a relatively small proportion of smokers who seek them out for support.[7] A more promising use of the telephone has been demonstrated.[8–10] Orleans and her colleagues[8] found that a combination of self-help materials and minimal telephone support resulted in 6-month continuous abstinence rates at 16 months that clearly favored telephone support over self-help materials alone.

An even more promising approach may be a combination of brief telephone support and outreach to identified smokers. Such an outreach strategy may be a cost-effective means of reaching large proportions of smokers. Lando et al.[11] were successful in recruiting 70% of an identified population of smokers to participate in a study evaluating the effectiveness of self-help quit materials. Overall outcomes were comparable to those obtained in a number of other evaluations of self-help in which participants were self-referred volunteers.[12–14]

In the present study we assessed brief telephone outreach as a recruitment and intervention tool. Intervention took place in Bloomington, Minn, one of the three education communities in the Minnesota Heart Health Program (MHHP), a 10-year research and demonstration project intended to reduce cardiovascular disease risk.[15] At the time of the present study, intervention in Bloomington had been going on for almost 5 years. The selection of Bloomington permitted relatively easy access to a large identified population of smokers.

Four discrete smoker populations were studied. The groups included current registrants in an extended community quit-smoking contest, previous MHHP smoking class or contest participants, and smokers who were identified at an initial heart health screening but who had never participated in MHHP smoking cessation activities. Over 60% of the overall Bloomington population and approximately 44% of Bloomington smokers had taken part in heart health screenings. The fourth group, Bloomington residents who had not attended this screening or any other MHHP activity, had been identified through an earlier telephone survey. Smokers in this survey who expressed interest in quitting were eligible for intervention.

The authors are with the Minnesota Heart Health Program at the Division of Epidemiology, School of Public Health, University of Minnesota, Minneapolis.

Address correspondence and requests for reprints to Harry A. Lando, PhD, Division of Epidemiology, School of Public Health, University of Minnesota, 1300 South Second Street, Suite 300, Minneapolis, MN 55454-1015.

This paper was submitted to the journal February 19, 1991, and accepted with revisions August 23, 1991.

TABLE 1—Number of Subjects Identified, Randomized, and Contacted for Follow-up

	Original Pool[a]	Randomized and Eligible for Follow-up	Contacted	
			6 mo.	18 mo.[b]
Quit Date '88				
Pledgers				
Intervention	156	129	126	117
Control	157	121	117	114
Nonpledgers				
Intervention	348	161	158	155
Control	198	143	135	130
Recyclers				
Intervention	448	215	207	200
Control	447	215	203	197
Heart health screenees				
Intervention	3462	215	201	189
Control	1100	218	201	196
Non-MHHP participants				
Intervention	241	217	204	194
Control	220	193	181	171
Total				
Intervention	4655	937	896	855
Control	2122	890	837	808

[a]Subjects who reported quitting prior to the intervention (n = 264) are included here as part of the original pool. However, these subjects were excluded from all analyses.
[b]Subjects who denied smoking more than 100 cigarettes in their lifetime (n = 16) at 6 months were not contacted at 18 months.

For each of the smoker populations, a supportive telephone outreach intervention was compared against a nonintervention control. In each case it was hypothesized that those assigned to telephone support would achieve a higher rate of quit attempts and long-term abstinence than would control subjects. It was further hypothesized that abstinence levels would be highest among current contest participants, followed by previous class or contest participants, and then by participants in screening who had not attended smoking cessation programs. No prediction of relative outcome was attempted for those who had not participated in MHHP activities or screening.

Methods

Subjects

Four mutually exclusive populations were included (Table 1).

Quit Date '83 registrants. The 859 subjects in this group had registered for an extended quit-smoking contest (Quit Date '88) that was offered in Bloomington from June 1988 through January 1989.[16-18] The subjects were further categorized by whether they pledged to quit on a specific date after entering the contest.

Recyclers. The 895 subjects in this group were identified from MHHP lists of former quit-smoking class or contest participants.

Heart health screenees. This group of 4562 subjects had identified themselves as smokers at heart health screenings up to 4 years earlier, but had not participated in MHHP smoking cessation activities. MHHP screening included assessment of the targeted risk factors: cholesterol, hypertension, obesity and eating habits, and smoking status.[15]

Nonparticipants in MHHP screenings and programs. A group of 410 smokers was identified from a total of 5642 telephone numbers. These numbers, randomly generated to represent a community population sample, did not appear on any list of MHHP participants. After these numbers were randomized to intervention and control conditions, a telephone survey was conducted. From this survey 4595 telephone numbers were identified as ineligible because they were incorrect or were business numbers or because no age-eligible smokers permanently resided in the household. A total of 1229 eligible smokers (between the ages of 25 and 74) were identified in 1047 households. Of these eligible smokers, 875 (71%) agreed to complete the telephone survey. Respondents were asked if they would be receptive to a subsequent telephone call that would offer them information on materials, classes, and methods to help them quit smoking. The 461 respondents (53% of those who completed the telephone survey; 37.5% of all eligible smokers identified) who expressed interest in this subsequent call were therefore eligible to participate in the study. Because of concern about the clustering of smoking behavior within households and the effects of subject interdependence on the analyses, only one smoker was randomly selected per household; thus 410 smokers were available for the study. (Randomization to intervention and follow-up.)

Of the total of 6777 smokers in these four populations, 4655 were randomized to intervention and 2122 to control. Among the 4655 smokers randomized to intervention, 1393 could not be reached. It is assumed that a majority of these were wrong numbers and thus ineligible. Smokers were also considered ineligible if they moved from Bloomington. Of the eligible 3262, only 65 (2%) refused to speak to the interventionist. A total of 813 persons had already quit smoking but remained in the intervention, which included a maintenance protocol. This left a pool of 2384 smokers. Those assigned to the control condition were not contacted prior to follow-up.

A smaller sample of 937 intervention and 890 control subjects was randomly selected for postintervention telephone follow-up at 6 and 18 months. At 6 months, 1733 (95%) of the 1827 subjects were contacted. At 18 months, follow-up was complete for 855 intervention and 808 control subjects, representing response rates of 91% for each group. Subjects who reported at the 6-month follow-up that they quit smoking prior to the intervention (n = 264) or who denied ever smoking more than 100 cigarettes in their lifetime (n = 16) were excluded from the analyses because the focus of the study was intervention effects on current smokers. Thus, 1453 (80%) of the 1827 subjects were classified as smokers at the time of the intervention and were included in the outcome analyses at 6 months; 1399 (77%) were available for the 18-month analysis. The 16 subjects who denied ever smoking more than 100 cigarettes, were not recontacted at 18 months.

Telephone Interventionists

Lay telephone interventionists included 15 women and 2 men ranging in age from 18 to 65 years. Ten interventionists had college degrees; seven had no college degrees but had previous telephone interviewing experience in the Division of Epidemiology.

Interventionists received 5 hours of training in smoking cessation techniques

and were also given individual reading assignments addressing health issues relating to smoking, benefits of smoking cessation, smoking cessation techniques, commonly asked questions pertaining to smoking, community resources dedicated to smoking cessation, and methods of supporting quit attempts. They were required to satisfactorily perform role plays of intervention calls and to make several calls to pilot subjects prior to telephoning actual study participants. A reference manual on cessation and maintenance strategies based on the MHHP *Quit and Win* self-help guide[19] was provided for their use in completing calls. A supervisor carefully monitored telephone calls, especially during the first few days of intervention. Weekly staff meetings were held to review progress and problems.

Procedure

Telephone calls were placed between 5 PM and 9 PM Sunday through Friday and between 9:30 AM and 1:30 PM on Saturday. Although a standard script was used for telephoning, the script allowed interventionists considerable latitude in talking to subjects. Interventionists introduced themselves and indicated that they were calling on behalf of the Bloomington Heart Health Program.

Current smokers were queried concerning their interest in quitting. Those who expressed interest were encouraged to initiate concrete steps toward cessation. The interventionist worked with the subject to identify possible barriers to cessation and methods by which these barriers could be minimized. Emphasis was placed upon the use of appropriate coping strategies to resist urges to smoke. Subjects who identified themselves as ex-smokers were congratulated and were asked if they had any questions about maintenance. They were offered a free maintenance guide and the telephone number of a smokers help line.

A major goal of the telephone intervention was to lead smokers to establish a specific target date for quitting. Subjects were offered free self-help materials based on the *Quit and Win* manual.[19] In addition, available help options were discussed, including Quit Date '88 (for individuals who were not already enrolled) and the community telephone help line service. Smokers were also given certificates that allowed them to enroll at no cost in a clinic program that was being offered to the general public for $95. Even if subjects were not prepared to set a quit date, they were urged to take concrete action toward quitting (e.g., cutting down on number of cigarettes, switching to a lower tar and nicotine brand).

All subjects were asked to consent to a second call within 2 to 3 weeks. For subjects who set a quit date, the second call was completed within 1 week after that date whenever possible. Subjects who had quit were congratulated and any difficulties were discussed. Those who had not initially set a date were again encouraged to do so. Subjects who had not quit or who were experiencing difficulties were reminded again of community resources. In both calls, the interventionist emphasized the identification, use, and modification of personally relevant coping strategies. First and second calls averaged less than 15 minutes each and were usually made by the same interventionist. Subjects were not offered the option of a third call. (A more detailed telephone intervention protocol is available on request from the first author.)

Follow-up Assessment

Telephone follow-up surveys at 6 and 18 months after the telephone intervention assessed basic demographic and smoking history data (at 6 months only), participation in cessation activities, information on quit attempts, and current smoking status. The surveys were administered by trained interviewers for the Telephone Survey Center in the Division of Epidemiology at the University of Minnesota. All subjects were asked, "Do you smoke cigarettes at present?" Those who said "no" at 6 months and who reported quitting after the intervention were scheduled for home visits to collect saliva samples for cotinine analysis. The mean interval between the telephone survey and saliva collection was 8 days; all samples were gathered within 15 days of the survey. No biochemical validations were performed at 18 months. Cotinine analyses were performed in the Division of Epidemiology.

Results

Initial Response to the Telephone Intervention

Of the 2449 smokers in the four intervention groups, 2384 (97%) received the intervention call and 75% requested the self-help smoking cessation materials offered by the interventionist. Seventy percent of the smokers consented to a second telephone call. Of those not already enrolled in Quit Date '88, 34.7% expressed interest in signing up for the contest in response to the telephone intervention. However, only 15% of the subjects committed to a specific quit date at the time of the initial telephone call. Among those who received a second call, 7.1% reported having quit, and an additional 14.6% reported having made a quit attempt since the first call.

Subject Characteristics

The 1399 subjects (intervention and control) available at 18 months included 703 women and 696 men. Demographic characteristics of the subjects are presented in Table 2. The mean age at which the subjects started smoking was 19 years. Their average age was 46.5 years. The subjects were asked about their current cigarette consumption or their cigarette consumption immediately prior to quitting. The average number of cigarettes smoked per day was 22. There were no significant differences between intervention and control subjects in demographic and smoking history characteristics. Generally, differences within the four subpopulations of smokers also were nonsignificant.

Quit attempts. Among current smokers, there were few differences in number of reported quit attempts (a minimum of 24 hours of abstinence) between intervention and control subjects or between the four smoker populations. This was true at both the 6-month follow-up and the 18-month follow-up.

Six-month abstinence. At the 6-month follow-up, a significant overall effect was found in favor of the intervention condition for both self-reported and cotinine-validated abstinence (Table 3). Overall self-reported abstinence rates at the time of the 6-month interview were 12% for intervention subjects and 8% for control subjects, $\chi^2(1) = 6.713$, $P = .009$. However, a substantial number of subjects (24%) refused to undergo cotinine validation. Of those who agreed, 45% exceeded the validation cutoff of 10 ng/mL. Saliva thiocyanate analyses were also available and were generally consistent with the results for cotinine. Thiocyanate levels exceeded 100 μmol/L (the recommended cutoff for validation) for 76% of those with cotinine levels greater than 10 ng/mL. Thus, validated 6-month abstinence rates were 5.48% for intervention subjects and 2.55% for control subjects, $\chi^2(1) = 8.018$, $P = .005$. Those who refused to undergo validation and individuals who were not contacted were classified as smokers for the purposes of data analysis.

TABLE 2—Demographic Characteristics of Combined Intervention and Control Groups Available at the 18-Month Follow-up

	Total, % (n = 1399[a])	Intervention, % (n = 716)	Control, % (n = 683)
Gender			
Female	50	50	50
Male	50	50	50
Age, years[b]			
0–24	1	1	1
25–34	19	19	18
35–44	28	28	27
45–54	25	25	25
55–74	28	27	29
Mean age (SEM)	46.49 (.47)	46.22 (.47)	46.78 (.48)
Employment			
Professional	17	16	18
Clerical	45	45	44
Blue collar	20	20	19
Not employed	19	18	19
Education			
High school or less	37	37	37
Post-high school	40	41	39
College graduate	23	22	25
Married	69	71	68

[a]Excludes subjects who denied ever smoking more than 100 cigarettes in their lifetime (n = 16) and those who quit prior to the intervention (n = 264).
[b]11 missing.

TABLE 3—Abstinent Subjects at 6- and 18-Month Follow-ups, by Group

	6 Months			18 Months	
	n[a]	Self-reported, %	Validated,[b] %	n[a]	Self-reported, %
Quit Date '88					
Pledgers					
Intervention	102	20	8	96	18
Control	103	13	4	100	21
95% CI (difference)		(−3, 17)	(−2, 10)		(−14, 8)
Nonpledgers					
Intervention	149	11	5	147	13
Control	128	9	4	124	18
95% CI (difference)		(−5, 9)	(−4, 6)		(−14, 4)
Recyclers					
Intervention	144	13	8	138	16
Control	141	6	1	137	16
95% CI (difference)		(0.2, 14)	(2, 12)		(−9, 9)
Heart health screeners					
Intervention	155	8	3	145	11
Control	155	7	1	155	11
95% CI (difference)		(−5, 7)	(−1, 5)		(−7, 7)
Non-MHHP participants					
Intervention	198	11	5	189	14
Control	178	7	3	169	14
95% CI (difference)		(−2, 10)	(−2, 6)		(−7, 7)
Total					
Intervention	748	12	5.5	716	14
Control	705	8	2.5	683	16
95% CI (difference)		(1, 7)	(1, 5)		(−6, 2)

[a]Excludes subjects who denied ever smoking more than 100 cigarettes in their lifetime (n = 16) and those who quit prior to the intervention (n = 264).
[b]Cotinine ≤10.

Although trends were in the expected direction within each of the four population groups, differences between the intervention and control groups were significant only for previous class or contest participants (recyclers). In this population, the self-reported abstinence rate at 6 months was 13.2% among intervention subjects, as opposed to only 5.7% for control subjects, $\chi^2(1) = 4.698$, $P = .03$. Again, however, validated abstinence rates were considerably lower: 7.6% for intervention subjects and 1.4% for control subjects, $\chi^2(1) = 6.33$, $P = .012$.

Eighteen-month abstinence. Differences between intervention and control subjects were no longer significant at the 18-month follow-up, either overall or for any of the separate smoker populations (see Table 3).

Continuous abstinence. Continuous abstinence was also examined at both the 6-month and 18-month follow-ups. Continuous abstinence was considered in three categories: validated at 6 months, self-reported at 6 months, and self-reported at 18 months. In all categories, the quit date had to be at least 3 months prior to the interview. Table 4 shows the number of subjects within each of these categories who reported "no puff" since quitting and no relapse lasting more than 6 days since quitting. Overall differences again favored the intervention condition at 6 months but not at 18 months. Differences were not significant within the four separate population groups. Overall, at the 6-month follow-up, validated 3-month continuous abstinence rates were 4.5% for intervention subjects and 2.1% for controls, $\chi^2(1) = 6.44$, $P = .01$. The most stringent criterion (6 months' continuous abstinence, not even a puff) yielded slightly lower abstinence rates: 3.8% for intervention subjects and 2.1% for controls, $\chi^2(1) = 3.64$, $P = .056$. As with point prevalence data smoking status at the time of follow-up, self-reported continuous abstinence rates were considerably higher than were cotinine-validated abstinence rates.

Consistent with the point prevalence data, overall 18-month self-reported continuous abstinence rates did not differ between intervention and control groups. However, when only that subset of subjects was considered who achieved at least 3 months of continuous self-reported and validated abstinence by the 6-month follow-up, differences at 18 months continued to favor the intervention condition.

Discussion

The initial results of the present study were moderately encouraging. Two brief telephone calls were sufficient to produce overall differences in outcome between intervention and control subjects at the

TABLE 4—Subjects Reporting Continuous Abstinence (>3 months), by Combined Intervention and Control Groups (n = 1399[a])

	Intervention (n = 716)		Control (n = 683)			
	n	%	n	%	$P(\chi^2)$	95% CI[b]
Validated at 6 months[c]	32	4.5	14	2.1	.011 (6.435)	0.54, 4.26
No puff at 6 months	27	3.8	14	2.1	.056 (3.64)	−.07, 3.46
No relapse > 6 days at 18 months	24	3.4	10	1.5	0.22 (5.25)	0.29, 3.51
No puff at 18 months	20	2.8	10	1.5	.086 (2.943)	−0.21, 2.81
Self-reported at 6 months	61	8.5	34	5.0	.008 (6.927)	0.88, 6.1
No puff at 6 months	43	6.0	24	3.5	.029 (4.759)	0.28, 4.72
No relapse > 6 days at 18 months	41	5.7	23	3.4	.035 (4.455)	0.13, 4.48
No puff at 18 months	31	4.3	18	2.6	.085 (2.969)	−0.21, 3.61
Self-reported at 18 months	84	11.7	84	12.3	.744 (0.106)	−4.0, 2.81
No relapse > 6 days	80	11.2	81	11.9	.688 (0.162)	−4.1, 2.65
No puff at 18 months	58	8.1	56	8.2	.946 (0.005)	−2.97, 2.77

[a]Excludes those who denied ever smoking more than 100 cigarettes in their lifetime (n = 16) and those who quit prior to the intervention (n = 264).
[b]CI = confidence interval for percentage difference between intervention and control groups.
[c]Twenty-seven percent of the subjects refused validation and were classified as smokers.

6-month follow-up. The success of this brief and potentially generalizable intervention in substantially increasing quitting rates among previous class or contest participants at that point is a positive finding. These results suggest that those who have enrolled in formal quit programs or contests may be receptive to further intervention despite the fact that they are likely to be hard-core smokers.[1] However, the overall cotinine-validated quit rate of less than 6% for intervention subjects at 6 months is modest.

The rather dramatic differences between self-reported and cotinine-validated abstinence rates were unexpected.[20,21] In addition to this high level of disconfirmation, more than one quarter of the subjects refused to undergo cotinine validation. All of these subjects were classified as smokers for the purposes of data analysis. This classification may be a bit conservative, however. Several of those who refused objected to the intrusiveness of the saliva collection procedure. Others expressed a fear of contracting AIDS as a result of providing a saliva sample. An increasing refusal rate for saliva collection has been noted elsewhere.[11]

Unfortunately, overall treatment effects were no longer evident at the 18-month follow-up. Somewhat encouraging is the fact that among subjects who had achieved at least 3 months of continuous abstinence by the 6-month follow-up, 18-month results continued to favor the intervention condition. Self-reported abstinence rates were higher overall at 18 months than at 6 months. This might have suggested a delayed impact of intervention,[13] except for the fact that abstinence rates actually tended to show greater improvement in control subjects over this time period. The 16% self-reported abstinence rate among control subjects at 18 months was surprisingly high, but self-reported quit rates are likely to have been inflated by false reporting. Number of quit attempts did not differ significantly between intervention and control subjects (although trends at the 6-month follow-up were in the expected direction).

Although the overall results were modest, the receptivity of smokers to the telephone intervention is a positive finding of this study. Previous studies have indicated that the vast majority of smokers are not interested in formal face-to-face interventions.[1,22] In the present study, 75% of intervention subjects accepted self-help materials and 70% agreed to a second telephone call. The level of responsiveness among the one group specifically recruited for the intervention—the non-MHHP participants—was especially gratifying. More than half of those interviewed not only expressed interest in quitting but agreed to receive intervention telephone calls. On the other hand, only 15% of all intervention subjects committed to a specific quit date at the time of the initial telephone call. Even so, respondents' acceptance of intervention offered through "cold" telephone contact suggests that telephone outreach procedures may be a useful channel for increasing the proportion of smokers recruited into treatment.

The demographic characteristics of the study participants support the ability of this intervention to reach individuals who are not typical participants in formal treatment programs. Over one third of all subjects reported a high school education or less. More than 40% worked in clerical positions and an additional 19% were not employed. In contrast to most formal programs, which tend to attract a more highly educated and professional clientele, the participants in this study were broadly representative of the larger Bloomington smoker population.

The cost of intervention could be relatively low if trained volunteers initiated telephone calls to identified populations of smokers. An organization such as the American Lung Association could provide additional telephone support and encouragement to individuals who had previously enrolled in its Freedom from Smoking clinics, for example. Communities could provide telephone support to smokers enrolled in local quit-smoking contests. Volunteer telephoners could be recent ex-smokers, whose support of others would reinforce their own efforts to remain abstinent. Telephone intervention is highly portable and could be used to reach individuals in remote locations.

The study population was drawn from a community that received multiyear heart disease prevention programming, thus raising questions about the study's generalizability. Furthermore, this community is located in Minnesota, a state that has been in the forefront of tobacco control activities for over a decade. These circumstances may have enhanced residents' responsiveness to any effort to change health behavior, including the current smoking cessation intervention. On the other hand, Bloomington residents may have experienced overexposure to MHHP intervention, and therefore may be resistant to undertaking further changes in their health behaviors.[18] National efforts supporting individual and

cultural changes in life-style–related education and social norms appear to be accelerating, especially with regard to smoking. The MHHP environment reflects what conceivably could become the norm in many parts of the nation.

The high degree of receptivity among the subjects and the positive initial results suggest that telephone intervention strategies are worth pursuing. At this point the challenge is to improve long-term maintenance. Future research should address the possible gain in efficacy versus the additional costs of more extended telephone support. If supportive telephone outreach strategies are replicable and can produce significant intervention effects, they may have important implications not only for smoking but for other life-style–related public health problems. □

Acknowledgment

This work was supported by a grant from the National Heart, Lung, and Blood Institute (1R01 HL25523).

References

1. Fiore MC, Novotny TE, Pierce JP, et al. Methods used to quit smoking in the United States: do cessation programs help? *JAMA.* 1990;263:2760–2765.
2. Schwartz JL. *Review and Evaluation of Smoking Cessation Methods: The United States and Canada 1978–1985.* Washington, DC: Department of Health and Human Services; 1987. NIH publication 87-2940.
3. Schachter S. Recividism and self-care of smoking and obesity. *Am Psychol.* 1982; 37:436–444.
4. Glynn TJ, Boyd GM, Gruman JC. Essential elements of self-help/minimal intervention strategies for smoking cessation. *Health Educ Q.* 1990;17:329–345.
5. Cohen S, Lichtenstein E, Prochaska J, et al. Some misconceptions about self-quitting: evidence from ten prospective studies of persons quitting smoking by themselves. *Am Psychol.* 1989;44:1355–1365.
6. Hatziandreu EJ, Pierce JP, Lefopoulou M, et al. Quitting smoking in the United States in 1986. *JNCI.* 1990;82:1402–1406.
7. Shiffman S, Read L, Maltese J, Rapkin D, Jarvik ME. Preventing relapse in ex-smokers: a self-management approach. In: Marlatt GA, Gordon JR, eds. *Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors.* New York: Guilford Press; 1985.
8. Orleans C, Schoenbach V, Wagner E, et al. Self-help quit smoking interventions: effects of self-help materials, social support instructions and telephone counseling. *J Consult Clin Psychol.* 1991;59:439–448.
9. Dubren R. Self-reinforcement by recorded telephone messages to maintain nonsmoking behavior. *J Consult Clin Psychol.* 1977;45:358–360.
10. Janis IC. The role of social support in adherence to stressful decisions. *Am Psychol.* 1983;38:143–160.
11. Lando HA, Pirie PL, McGovern PG, Pechacek TF, Swim J, Loken B. A comparison of self-help approaches to smoking cessation. *Addict Behav.* In press.
12. Cummings KM, Emont SL, Jaen CK, Sciandra R. Format and quitting instructions as factors influencing the impact of a self-administered quit smoking program. *Health Educ Q.* 1988;15:199–216.
13. Davis AL, Faust R, Ordentlich M. Self-help smoking cessation and maintenance programs: a comparative study with 12-month follow-up by the American Lung Association. *Am J Public Health.* 1984; 74:1212–1217.
14. Prochaska JO, DiClemente CC. Common processes of self-change in smoking, weight control, and psychological distress. In: Shiffman S, Wills TA, eds. *Coping and Substance Use.* Orlando, Fla: Academic Press; 1985:367–386.
15. Mittelmark MB, Luepker RV, Jacobs DR, et al. Community wide prevention of cardiovascular disease: education strategies of the Minnesota Heart Health Program. *Prev Med.* 1986;15:1–17.
16. Glasgow RE, Klesges RC, Mizes JS, Pechacek TF. Quitting smoking: strategies used and variables associated with success in a stop-smoking contest. *J Consult Clin Psychol.* 1985;53:905–912.
17. Lando HA, Loken B, Howard-Pitney B, Pechacek T. Community impact of a localized smoking cessation contest. *Am J Public Health.* 1990;80:601–603.
18. Lando HA, Hellerstedt W, Pirie P, Fruetel J, Huttner P. Results of a long-term community smoking cessation contest. *Am J Health Promotion.* 1991;5:420–425.
19. Pechacek T, Arkin R, Johnston-Hall N. *Quit and Win.* Minneapolis, Minn: University of Minnesota, Minnesota Heart Health Program; 1984.
20. Fortmann S, Rogers T, Vranizan K, Haskell WL, Solomon DS, Farquhar JW. Indirect measures of cigarette use: expired-air carbon monoxide versus plasma thiocyanate. *Prev Med.* 1984;13:127–135.
21. Stookey GK, Katz BP, Olson BL, Drook CA, Cohen SJ. Evaluation of biochemical validation measures in determination of smoking status. *J Dent Res.* 1987;66:1597–1601.
22. Jeffery RW, Hellerstedt WL, Schmid TL. Correspondence programs for smoking cessation and weight control: a comparison of two strategies in the Minnesota Heart Health Program. *Health Psychol.* 1990; 9:585–598.

Adolescent Smoking, Weight Changes, and Binge-Purge Behavior: Associations with Secondary Amenorrhea

Jim Johnson, PhD, and Agnes H. Whitaker, MD

ABSTRACT

Background. The association of secondary amenorrhea with extreme forms of substance use, weight control, and exercise in nonrepresentative samples raises questions as to whether adolescents in the general population who engage in these behaviors are at increased risk for secondary amenorrhea. We examined the prevalence and behavioral correlates of secondary amenorrhea in a county-wide high school population of 2544 girls aged 13 to 18.

Methods. A survey questionnaire, which elicited menstrual history as well as weight history, weight control practices, level of exercise, and use of cigarettes, wine, and beer, was administered during school hours; absentees were also surveyed. The completion rate was 91%.

Results. The 1-year prevalence of secondary amenorrhea was 8.5%. Secondary amenorrhea was associated with smoking one or more packs of cigarettes per day (adjusted relative risk [RRa] = 1.96, 1.21–3.10), with multiple binge-eating behaviors in combination with laxative use or self-induced vomiting (RRa = 4.17, 2.54–6.32), and with weight fluctuation due to weight control (RRa = 2.59, 1.33–4.79). There was no association between amenorrhea and alcohol consumption or exercise level.

Conclusions. Estimates of attributable risk are provided and indicate that bulimic behaviors and cigarette smoking may result in a considerable excess of cases of secondary amenorrhea in an adolescent population. (*Am J Public Health.* 1992;82:47–54)

Introduction

There is a striking paucity of data on the epidemiology and correlates of amenorrhea among adolescent females. The few studies that have examined menstruation among community-based samples of adolescent females have restricted their focus to age at menarche[1–4] and to other menstrual characteristics, such as oligomenorrhea[1] and dysmenorrhea.[1,5] In recent years, clinical research on amenorrhea has also focused on nonpregnant clinical samples with high rates of amenorrhea. In addition to patients with well-defined medical disorders,[6,7] these samples include patients with eating disorders, such as anorexia nervosa[8–10] and bulimia,[11–16] and substance abuse disorders, such as alcoholism.[17] High rates of amenorrhea have also been found in career ballet dancers[18,19] and athletes.[20–22]

The association of amenorrhea with extreme forms of weight control, substance use, and exercise in these highly nonrepresentative samples raises the question of whether adolescents in the general population who engage in these behaviors are at increased risk for secondary amenorrhea. The importance of the question is underscored by two sets of findings: (1) the high rates at which adolescent girls experiment with weight control,[23] cigarettes,[24,25] and alcohol;[25,26] and (2) the association of amenorrhea in selected samples with increased rates of infertility,[27] increased rates of osteoporosis,[28,29] scoliosis and stress fractures,[30,31] and lower lifetime occurrence of breast cancer and cancers of the reproductive system[32]—all conditions of major concern to women.

Amenorrhea is classified for clinical purposes as either primary or secondary. Primary amenorrhea is defined as the failure to achieve menarche by age 16.[6,7] Secondary amenorrhea refers to the prolonged absence (usually defined as 3 months) of menstrual bleeding any time after menarche.[6,7] These age and duration guidelines for the clinical investigation of adolescent amenorrhea take into account the normal variation in age of menarche (11 to 15 years)[2] and the physiological occurrence of irregular intermenstrual intervals (oligomenorrhea) due to anovulatory cycles in the first years after menarche.[1,33–35] Amenorrhea is properly regarded as a symptom that may reflect quite different etiologies and endocrine profiles.

This study focuses on the prevalence and behavioral correlates of secondary amenorrhea in an unselected population of adolescent females.

Methods

Study Sample and Measures

The study sample was a county-wide population of high school students (n = 5596) who participated in a survey of selected psychiatric disorders in 1985.[23] Students were surveyed during classroom hours in the eight high schools in the county (five general public schools, one vocational public school, one independent

Jim Johnson is with the Clinical and Genetic Epidemiology Unit of Columbia University College of Physicians and Surgeons. Agnes H. Whitaker is with the New York State Psychiatric Institute, Division of Child Psychiatry at the Columbia University College of Physicians and Surgeons.

Requests for reprints should be sent to Agnes H. Whitaker, MD, Box 78, New York State Psychiatric Institute, 722 West 168th Street, New York, NY 10032.

This paper was submitted to the journal September 24, 1990, and accepted with revisions May 7, 1991.

TABLE 1—Unadjusted Relative Risks for Secondary Amenorrhea

Risk Factor	Comparison	% Exposed	Relative Risk	95% CI
Social class (I, II)	III-V	19.2	1.27	.89–1.80
Age 13	ages 15–17	1.5	1.42	.53–3.32
Age 14	ages 15–17	19.4	0.99	.69–1.40
Age 18	ages 15–17	4.0	1.26	.92–2.25
Age at menarche >12.5 years	younger than 12.5 years	51.3	1.48	1.12–1.93
First 2 years of menarche	post first 2 years	48.1	1.62	1.15–2.24
BMI lowest decile	middle 80%	10.7	1.05	.66–1.64
BMI highest decile	middle 80%	10.4	1.11	.72–1.69
Diet ≤ 28 days	nondieters	43.5	0.92	.68–1.24
Diet > 28 days	nondieters	18.5	0.91	.61–1.33
Fast ≤ 2 days	nonfasters	12.7	1.39	.96–1.99
Fast ≥ 3 days	nonfasters	7.4	1.84	1.20–2.71
Frequent self-induced vomiting	nonvomiters	4.0	2.58	1.61–3.94
Infrequent self-induced vomiting	nonvomiters	11.7	2.09	1.49–2.85
Frequent laxative use	no use	1.6	3.31	1.78–5.44
Infrequent laxative use	no use	3.8	1.42	.76–2.50
Diet pills, frequent	no use	5.1	1.84	1.10–2.94
Diet pills, infrequent	no use	15.7	1.48	1.06–2.04
Frequent binge eating	no binge	10.2	2.17	1.32–3.44
Infrequent binge eating	no binge	70.6	1.53	1.01–2.26
Three meals per day	other pattern	37.4	0.71	0.52–0.96
Weight gain ≥ 4.5 kg	no change ≥ 4.5 kg	22.2	1.93	1.36–2.67
Weight loss ≥ 4.5 kg	no change ≥ 4.5 kg	10.3	1.54	1.05–2.23
Both loss and gain	no change ≥ 4.5 kg	3.5	2.82	1.61–4.60
Exercise every day	twice or less/wk	14.6	0.83	0.53–1.28
Exercise 3–6 times/wk	twice or less/wk	52.2	0.83	0.64–1.12
Alcohol (≥twice/wk)	no use	8.8	1.75	1.15–2.58
(several times/mo)	no use	31.1	1.19	0.88–1.60
Smoking ≥ one pack/day	nonsmokers	6.8	2.23	1.48–3.26
Smoking < one pack/day	nonsmokers	25.3	1.31	0.96–1.76

Note. CI = confidence interval.

day and boarding school, and one parochial day school). The study protocol included provision for surveying absentees. Of the county-wide enrollment, 91% (2564 boys and 2544 girls) returned valid questionnaires. The population was primarily White (94 percent) and was evenly distributed across ages 14 to 17; 2.5% were 13 years old or younger, and 5.5% were 18 years old or older.

The girls' version of the survey questionnaire asked whether menstrual periods had begun and, if so, the year and month of menarche. Postmenarchal girls were asked, "Have you missed three of your expected periods in a row since this time last year?" Other menstrual questions, with the exception of the item concerning amenorrhea, were patterned after the questionnaire items of Widholm and Kantero.[1] Other investigators have found that adolescents report age at menarche with a high degree of accuracy.[36]

Weight control practices, binge-eating, and weight history were assessed with the Eating Symptoms Inventory.[23] This instrument contains 24 sets of questions that operationalize the symptoms of anorexia nervosa and bulimia as defined by the American Psychiatric Association in *the Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.) (*DSM-III*).[37] Binge-eating behaviors included large binges, recurrent binges, and binge discomfort (termination of binges for abdominal pain, sleep, social interruption, or feeling the urge to throw up), as detailed in an earlier report.[23] Dieting, fasting, purging (self-induced vomiting and laxative use), and use of diet pills and diuretics were also assessed. From a review of semistructured clinical interviews with a subsample of 356 students who participated in the survey, a screen for eating disorders based on items in the inventory and on the Diet Factor from the Eating Attitudes Test[38] identified cases of anorexia nervosa and bulimia with a sensitivity of 86.7 and a specificity of 72.9.[39]

Level of exercise was assessed by listing nine types of activities and asking how many days per week in the past month the student engaged in those activities. Cigarette use and beer and wine consumption were assessed with items drawn from Kandel.[40] Other studies have found that adolescent self-reports on cigarette[41] and alcohol use are valid.[42,43] To our knowledge, the validity of adolescent reports on exercise has not been studied.

As for background characteristics, social class was defined according to Hollingshead,[44] relative body weight was defined by body mass index (BMI) (weight in kg/height in $cm^2 \times 10\,000$),[45,46] and gynecologic age was defined as the difference in years between chronological age and age at menarche.[47,48] Accuracy of students' reports of height and weight were assessed with a validity study.[23] Reports for height were found to be accurate and unbiased whereas heavy girls tended to underestimate their weight; however, the bias was minimal (~2.5 kg).

By definition, premenarchal girls were not considered to be at risk for secondary amenorrhea.

Statistical Analysis

Risks are framed in terms of relative risk (RR).[49] Problems related to interpretation of risk factors in a cross-sectional study are reserved for the discussion. We also report 95% confidence intervals (CIs) based on the limits of risk.

Most of the putative risk factors (weight history, weight control practices, substance abuse, and exercise) were to some extent correlated; any statement about independent contributions to the risk of secondary amenorrhea had to include evaluation of these potential risk factors in a multivariate logistic regression model in conjunction with careful examination of the cross-classified risks. The RRs of amenorrhea reported in Tables 1 and 2 are the ratios (proportions) of the risk in those exposed to the risk factor to the risk in those unexposed. CIs are from the limits of the probabilities of a logistic regression estimate in both the unadjusted (Table 1) and adjusted (Table 2) case. In Table 2, the estimated risk for an individual is given by

$$1/\{1 + \exp[-1*(\text{constant} + b1X1 + \ldots brXr)]\}.$$[50]

TABLE 2—Summary from Multivariate Analyses (Logistic Regression): Risk Factors for Secondary Amenorrhea among Nonreferred Adolescent Girls

Risk Factor	Exposure	Adjusted Relative Risk	95% CI	Excess Fraction for Population	Attributable Risk in Those Exposed
Smoking (≥one pack daily)	.068	1.96	(1.21–3.10)	.061	.49
(<one pack daily)	.253	1.35	(0.90–1.72)	.059	.20
Binge-purging (frequent)	.031	4.17	(2.54–6.52)	.089	.76
(infrequent)	.143	1.99	(1.39–2.81)	.124	.50
Weight loss (≥ 4.5 kg)	.103	1.45	(.95–2.20)	.044	.31
Weight gain (≥ 4.5 kg)	.222	1.71	(1.16–2.49)	.136	.42
Both loss and gain	.035	2.59	(1.33–4.79)	.053	.61
First year after menarche	.129	1.74	(1.07–2.76)	not applicable	
Age at menarche > 12.5 years	.481	1.38	(0.99–1.92)	not applicable	

Note. CI = confidence interval.

We have avoided reporting adjusted odds ratios in this cross-sectional study, given that the risk in some of the exposed strata (e.g., severe binge-purge) violates the "rare disease" assumption.[49] If one is concerned about statistical significance, the CIs reported in Tables 1 and 2 exclude one only when the odds ratio would also. To the extent that the risk in a number of the exposed groups exceeds 10%, the relative risks, as reported in Tables 1 and 2, is the superior point estimates of risk when compared with the odds ratio.

As an estimate of the adjusted population excess fraction (population attributable risk), we also use

$$p(\text{exposure})^*(1 - RRa)/$$
$$\{[p(\text{exposure})^*(RRa - 1)] + 1\}.$$

Because there is some controversy over the use of this estimate,[51,52,53] we stratified the data as far as feasible and reestimated using (1) Greenland's "exact method"[52] equation 2, and (2) Miettinen's multivariate confounder score method; the latter allowed us to do what we did in (1) while retaining more information with respect to dosage.[54] In each case, the estimates obtained by these three alternative methods were consistent within rounding error. Convergence of solutions suggests that the estimates of adjusted excess fractions (Table 2) are reasonable.

Results

Prevalence of Primary and Secondary Amenorrhea. Among girls aged 16 and older, the prevalence of primary amenorrhea (premenarchal status) was 2.0%. The 1-year prevalence of secondary amenorrhea in postmenarcheal girls was 8.5%. Only 1 out of the 245 girls subsequently interviewed by clinicians in the second stage of the study had ever been pregnant, suggesting that pregnancy at most accounted for one or two reports of secondary amenorrhea in these data. Among the 2544 girls surveyed in this semirural county, 36 reported using birth control pills, and the prevalence of secondary amenorrhea among them was 5.6% (vs 8.5% for the remainder of the population).

Relationships with Social Class, Age, and BMI. The proportion of girls who reported secondary amenorrhea in the past year appears in Table 3 by social class,[44] chronological age, gynecological age, and BMI quartile. Secondary amenorrhea was not significantly related to social class (by chi-square statistic). A post hoc grouping of the lowest two social classes (suggested by the data in Table 3) also failed to yield a significant difference. Risk of secondary amenorrhea was highest among girls whose gynecologic age was 2 years or less, and it was lower in the older gynecologic age groups.

There was no relationship between secondary amenorrhea and BMI quartile. More restricted divisions of BMI, including deciles, were also unrelated to secondary amenorrhea, as was percent below ideal weight at lowest weight for previous year.

Weight Control, Bulimic Behaviors, and Weight Changes. Girls who had dieted, even for longer than 28 consecutive days, were not at increased risk for secondary amenorrhea when compared with girls who had not. However, girls who had fasted for extended periods (3 or more days) were at some increased risk (RR = 1.84, 95% CI = 1.20–2.71) (Table 1). Duration of dieting and fasting were evaluated independently of kilograms lost on a diet or fast because many girls reported that they had dieted (or fasted) for extended periods without weight loss.[23]

In contrast to simple duration of dieting, other methods of weight control

TABLE 3—Rate/100 of Adolescent Females Who Reported Missing Three Consecutive Menstrual Periods in the Past Year, by Background Characteristics (n = 2156)

	Rate/100 Secondary Amenorrhea
Socioeconomic status (SES) quintiles	
I (low SES)	10.4
II	10.2
III	7.5
IV	8.5
V (high SES)	8.5
Body mass index (BMI) quartile	
I (thin girls)	8.9
II	9.3
III	8.1
IV (heavy girls)	7.8
Chronological age	
13	10.8
14	7.8
15	9.4
16	7.9
17	8.1
18	9.8
Gynecologic age	
0	12.5
1	13.5
2	9.7
3	9.6
4	6.1
5	8.5
6	5.3
7	5.4

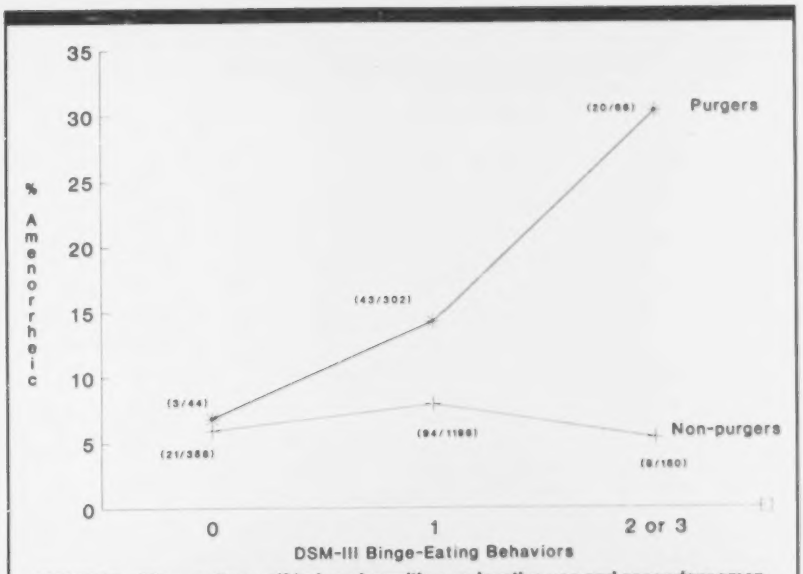


FIGURE 1—Binge eating, self-induced vomiting, or laxative use and secondary amenorrhea. Amenorrhea is defined as three consecutive missed menstrual periods in the past year in a postmenarchal girl. "Purgers" report use of self-induced vomiting or laxative for weight control. Binge-eating behaviors consist of large binges, frequent binges, and binge discomfort (see text).

were associated with higher risk for secondary amenorrhea. Self-induced vomiting, use of laxatives, and use of diet pills (mainly nonprescription) were each associated with one and one-half to threefold increases in risk for secondary amenorrhea (Table 1). The use of diuretics (nonprescription) was not significantly associated with increased risk (RR = 1.64, 95% CI = n.s.); however, the number of girls who used diuretics was small (n = 22), and thus our statistical power to detect such an increase in risk was minimal. As we have reported elsewhere,[23] weight control behaviors were more common in girls who were older and heavier (by BMI) and were minimally related to social class.

Reports of binge-eating consistent with the *DSM-III*[37] definition of this behavior were associated with a twofold increase in the risk of secondary amenorrhea.

With respect to weight change in the past year, girls who had gained at least 4.5 kg, or 10 lb, and girls whose weight had fluctuated (a gain of 4.5 kg or more combined with a loss of 4.5 kg or more) in the past year were at higher risk for secondary amenorrhea (RR = 1.93, 95% CI = 1.36–2.67; RR = 2.82, 95% CI = 1.61–4.60, respectively) when compared with girls who had not gained or lost 4.5 kg in the previous year (Table 1). Girls who reported a simple loss of at least 4.5 kg in the previous year were also at some increased risk for secondary amenorrhea (RR = 1.54, 95% CI = 1.05–2.23). However, a weight loss of 4.5 kg or more appeared to be a greater risk factor in girls in the lowest BMI quartile (RR = 3.00, 95% CI = 1.44–5.53, not shown on Table 1). For these analyses, a cutoff of 4.5 kg or more was used to evaluate weight change as a risk factor. Parallel analyses using number of kilograms lost, percent of body weight lost, or various transformations (e.g., log (kg)) produced similar results.

Exercise Regimens. Girls who exercised daily were not at greater risk for premenarchal status or secondary amenorrhea than those who did not. Gymnasts and dancers who exercised daily also failed to show an elevated risk (RR = 0.99, 95% CI = 0.54–1.77, not shown on Table 1).

Cigarette Smoking and Alcohol Use. The risks associated with cigarette smoking and alcohol also appear on Table 1. Both risk factors show evidence of a dose-response relationship with twofold increases in secondary amenorrhea at the higher dose (one pack of cigarettes or more per day; alcohol use more than several times per week). Cigarette smoking and alcohol use in this sample were related: girls who smoked were twice as likely to drink, and girls who drank were twice as likely to smoke. Both variables were related to age, with older girls more likely to both smoke and drink. Neither variable was related to social class. In contrast to adults, smoking here was not related to BMI: smokers in this adolescent sample were *not* of lighter relative weight than nonsmokers.

Risk Factors for Secondary Amenorrhea: Multivariate Analysis. Table 2 presents the final model from successive logistic regression and cross-classification analyses.[55–56] It reports the probability of exposure (percent of girls with the characteristic), the RRa, the excess fraction for the population, and the attributable risk in those exposed to a risk factor.[50]

Of particular importance to clinicians is binge-purge behavior. As Figure 1 shows, the risk associated with binge-eating behavior was modified by the level of purging (laxative use or self-induced vomiting). The statistical interaction suggested by Figure 1 was significant ($P < .001$). Because laxative use in combination with binge-eating rarely occurred in the absence of self-induced vomiting (n = 69), laxative use was combined with self-induced vomiting for the purposes of these analyses. Binge eating was a risk factor for secondary amenorrhea only when it occurred in combination with purging (and vice versa). Given the striking pattern of results in Figure 1, binge-purge behavior was modeled with two variables: partial binge-purge behavior, defined as self-induced vomiting or laxative use in conjunction with a single *DSM-III* binge-eating behavior (RRa = 1.99, 95% CI = 1.39–2.81), and complete binge-purge behavior, defined as self-induced vomiting or laxative use in conjunction with two or three *DSM-III* binge-eating behaviors (RRa = 4.17, 95% CI = 2.54–6.52) (Table 2).

Girls who smoked a pack of cigarettes or more per day were at excess risk for secondary amenorrhea, even with statistical control for other variables in the model. The variable representing lower levels of exposure to cigarettes (less than a pack per day) was not statistically significant in the final multivariate model, but the point estimate of the effect (RRa = 1.35, 95% CI = 0.90–1.72) (Table 2) was intermediate, consistent with a dose-response relationship. To the extent that there were age differences in the association between smoking and secondary amenorrhea, there was a nonsignificant tendency for the association to be stronger in older girls. Unfortunately, no information was available as to how long the girls had been smoking.

A number of variables contributed little to the model after adjustment for other associated variables. To the extent that diet pills, diuretics, or simple duration of dieting or fasting was associated with increased rates of amenorrhea, it appeared to be due to the association of these variables with other risk factors (i.e., weight fluctuation and binge-purge behavior).

The decision to drop alcohol from the final model was made only after examining the effects of alcohol consumption across levels of other risk factors. No excess risk was apparent above and beyond the effects of confounding variables.

Focusing on the population excess fractions in the fourth column of Table 2, it is apparent that many of the cases of secondary amenorrhea in the population were associated with behavior: 12% with smoking and 21% with binge-purge behavior. The direction of effect was less compelling with respect to the fraction associated with weight gain, loss, and fluctuation. However, most weight fluctuation in the population coincided with some form of weight control behavior (85%). The final column of Table 2 provides adjusted estimates of attributable risk in girls actually exposed to the model risk factors.

It was not possible to confirm the results of the model presented in Table 2 with a complete cross-classification of risk factors because the number of possible combinations is unwieldy (8640, if BMI is categorized). As an alternative, the data of primary interest are presented in Figure 2, cross-classifying and collapsing dose categories across smoking, binge-purge behavior, and type of weight change.

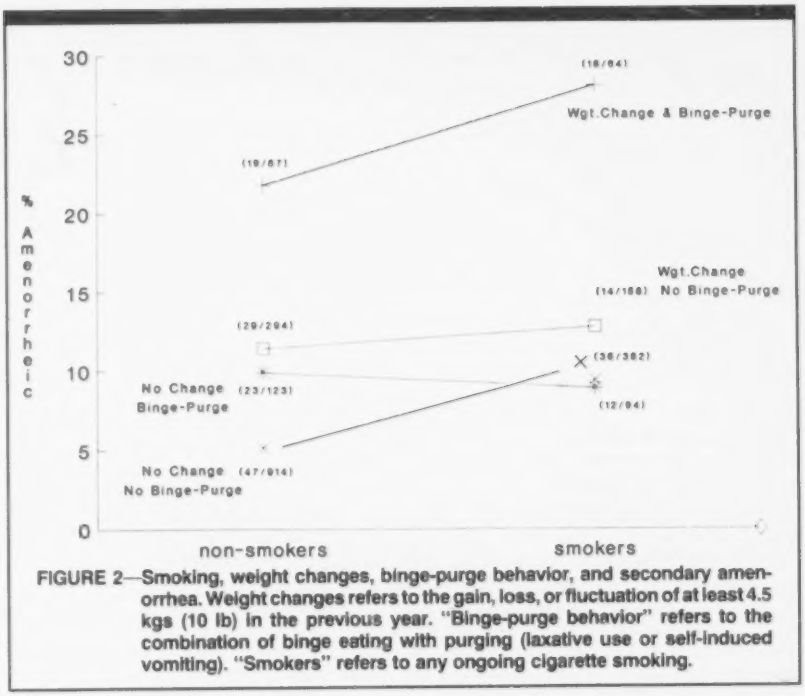


FIGURE 2—Smoking, weight changes, binge-purge behavior, and secondary amenorrhea. Weight changes refers to the gain, loss, or fluctuation of at least 4.5 kgs (10 lb) in the previous year. "Binge-purge behavior" refers to the combination of binge eating with purging (laxative use or self-induced vomiting). "Smokers" refers to any ongoing cigarette smoking.

Discussion

The prevalence of secondary amenorrhea in an unselected adolescent population has not been previously reported. The most methodologically satisfactory study of secondary amenorrhea among adult women (excluding instances attributable to pregnancy and surgical treatment) found a 1-year prevalence of 4.4% for women aged 18 to 45, with a 6.3% prevalence in the age group 18 to 24[57]—somewhat lower than the rate of 8.5% found in this study. Lower rates among adult women may be due to a number of factors, such as maturation of the reproductive cycle[47] or lower rates of the types of behavior examined here.

Certainly, secondary amenorrhea may be due to other underlying medical causes, which must be systematically excluded in a thorough diagnostic workup of each case.[6,7,58] However, such unmeasured risk factors would have to be strongly associated with the risk factors identified by this observational study to render these results spurious. We find an explanation pointing to spurious associations unlikely, given the biologic plausibility of the risk factors nominated.

Binge-eating behaviors were associated with a markedly increased risk for secondary amenorrhea. An association between binge-purge behavior and amenorrhea and/or menstrual irregularity among high school girls was noted by Johnson et al.,[59] but rates and risks for amenorrhea were not fully explored. With respect to the risk for secondary amenorrhea associated with binge-purge behavior, it is likely that purging (self-induced vomiting or laxative use) rather than binge eating is the causal agent. Little, if any, increased risk was evident in the large number of girls who binged but did not purge. The fact that the combination of binge eating and purging carries increased risk is likely to indicate the (unmeasured) increased frequency of purging accompanying the full binge-purge behavioral syndrome.

The mechanism by which binge-purge behavior may affect the hypothalamic-pituitary-ovarian axis is not clear. At least three separate studies[60–62] find evidence in adult bulimic patients of disturbed hypothalamic-pituitary function that is not attributable to emaciation. Possible etiologic factors include nutritional deficiencies,[63,64] insufficient body fat content,[65] intermittent dieting,[66] stress,[60,67,68] exercise,[21] and adaptation to semistarvation at a weight that is considered "normal" by population standards but is below the physiological setpoint for the individual.[62]

Weight changes resulting from weight control practices were associated with secondary amenorrhea above and beyond the specific methods used (Table 2). In particular, weight fluctuation over the past year was associated with the greatest risk for secondary amenorrhea.

This epidemiologic study confirms the importance of weight control in the pathogenesis of amenorrhea in "normal weight" women, a phenomenon long noted in gynecologic case reports[69,70] and recently cited as an underrecognized contributor to infertility.[71] The episodic loss and regain of weight that is so common among adolescent girls[23] may be far from benign.

In these data, the excess risk associated with weight fluctuation in conjunction with binge-purge behaviors may be a proxy for the most severe forms of binge purging or for extremely rapid weight loss. We have controlled statistically for significant weight loss (4.5 kg) in the past year,

but these data did not allow us to take into consideration how rapidly these weight changes occurred.

Several clinical studies have also found a relationship between weight fluctuation and amenorrhea. Copeland and Herzog[16] found that, among normal weight bulimic patients without a history of anorexia nervosa, those who had gained weight from a previously lower weight had a much higher occurrence of amenorrhea than those who had not. Kreipe et al.[72] found a similar association between weight fluctuation and menstrual irregularity in women with subclinical eating disorders. They suggested two possible explanations: (1) these women have a natural weight setpoint that is higher than the weight considered normal on standardized tables; and (2) these women embark on unrealistic "crash" diets in which rapid weight loss (which is only to be regained) disrupts the menstrual cycle. A third possibility, proposed by Devlin et al.,[62] is that weight fluctuation per se may influence hypothalamic-pituitary-ovarian axis functioning.

The risk associated with a simple gain of 4.5 kg or more may reflect weight "rebound" in girls who attempted weight control. This possibility is suggested by the observation that a gain of 4.5 kg was more than would be expected from normal growth tables for girls aged 13 to 18. In addition, girls who had gained weight but had not engaged in any weight control behaviors were at a low risk for secondary amenorrhea (5/103 = 4.9%). A simple weight loss of 4.5 kg or more in the absence of regain or binge-purge behavior was not a strong risk factor for secondary amenorrhea, except among girls at very low relative body weight.

In this adolescent population, cigarette smoking was associated with secondary amenorrhea. Previous studies in adults have found cigarette smoking to be associated with higher rates of physician-attended menstrual disorders,[73] infertility,[74] low birth weight in offspring,[75] spontaneous abortion,[76] and early menopause.[77,78] Several biologic mechanisms have been proposed, such as actions on the hypothalamic-pituitary-ovarian axis, a direct toxic effect on the ovary, and alteration of peripheral estrogen.[79]

The results reported here suggest that as much as 50% of cases of secondary amenorrhea among adolescent girls may be related to behaviors that could be modified. Measures of attributable risk should not be interpreted as implying causality in the absence of other evidence for causality.[50] Nonetheless, these estimates, if cautiously interpreted, can identify the risk groups within a population. Adjusted estimates are also informative even if simple additivity of individual estimates is fallacious. We have interpreted these associations as risks from cross-sectional data on the basis of consistency with clinical studies and biologic plausibility.

Finally, a number of caveats are in order. First, it must be emphasized that underlying medical conditions could have been responsible for both weight fluctuations and the comorbid amenorrhea. A medical workup may thus be warranted, depending on the age of the adolescent and other factors.[6,7] At the population level, however, 85% of the adolescent girls who report weight changes greater than 4.5 kg also engage in some form of weight control, suggesting that the workup should include a careful review of weight control practices as well as smoking.

Second, the failure to find an association between exercise and amenorrhea could be a question of exposure level. Although we examined the risk of secondary amenorrhea in the most extreme exercisers in this population, it is likely that they are not comparable to the career dancers and athletes who have often been the focus of other reports.[18,19,22] However, Wilson et al.[80] also found no relation between exercise levels and menstrual characteristics in healthy adolescent girls. In addition, a number of reports have suggested that a significant amount of amenorrhea in athletes is related to binging and purging.[19,81]

Third, in contrast to clinical studies of menstrual disturbance in women with alcoholism,[17] these data did not implicate alcohol as a risk factor after adjustment for confounding variables (e.g., smoking). There are at least two alternative explanations for this lack of association in these data: (1) the exposure to alcohol was not severe enough or long enough to induce secondary amenorrhea in this young population, or (2) purging and smoking were measured more accurately than alcohol consumption in this study. Because either scenario would produce no association, alcohol cannot be exonerated as a risk factor for amenorrhea.

Fourth, these data are from self-report questionnaires, which doubtless have limitations. However, other investigators have validated adolescent self-reports of menstruation.[36] Our own clinical follow-up of 245 of these girls[39] suggests excellent agreement between the questionnaire data and the clinical interview with respect to both amenorrhea and high-risk eating behaviors. Furthermore, any misreporting of the presence of either secondary amenorrhea or high-risk behaviors would most likely bias against finding associations.

In summary, girls who engaged in binge-purge behavior were at risk for secondary amenorrhea. Second, weight changes, especially weight fluctuation, also appeared to disrupt menstruation. Finally, cigarette smoking was associated with secondary amenorrhea. These data suggest that evaluation of the adolescent with secondary amenorrhea, in both clinical and research settings, should include a careful history to assess these potentially modifiable contributory factors. □

Acknowledgments

This research was supported in part by National Institute of Mental Health (NIMH) grant 1–RO3–MH40527–01 ("Eating Disorders in Adolescents: An Epidemiologic Study") and Biomedical Research Support Grant (New York State Psychiatric Institute); NIMH grant ST32 MH 16434 and NIMH Clinical Research Center grant 30906–07. We thank David Shaffer FRCP (Lond), FRC Psych (Lond), Mark Davies, MPH, and Helen Myers for their invaluable earlier contributions to the study presented here; and Bruce Link, PhD, Diane E. McLean, PhD, Richard Neugebauer, PhD, Steve Ng, MD, Mervyn Susser, MB, BCh, FRCP(E), DPH, and B. Timothy Walsh, MD, for their helpful suggestions with the manuscript.

References

1. Widholm O, Kantero RL. A statistical analysis of the menstrual patterns of 8000 Finnish girls and their mothers. *Acta Obstet Gynecol Scand.* 1971;14(suppl):1–36.
2. National Center for Health Statistics. Age of menarche, United States. *Vital and Health Statistics.* Series 11. No. 33. Health Services and Mental Health Administration. Washington, DC: US Government Printing Office; 1973:1–29.
3. Zacharias L, Rand WM, Wurtman RJ. A prospective study of sexual development and growth in American girls: the statistics of menarche. *Obstet Gynecol Survey.* 1976;31:325–337.
4. World Health Organization. World Health Organization multicenter study on menstrual and ovulatory patterns in adolescent girls: I. a multicenter cross-sectional study of menarche. *J Adolesc Health Care.* 1986;7:229–235.
5. Klein JR, Litt IF. Epidemiology of adolescent dysmenorrhea. *Pediatr.* 1981;68:661–664.
6. Mansfield MJ, Emans SJ. Adolescent menstrual irregularity. *J Reprod Med.* 1984;29:399–410.
7. Soules MR. Adolescent amenorrhea. *Pediatr Clin North Am.* 1987;34:1083–1103.
8. Warren MP, Vande Wiele RL. Clinical and

metabolic features of anorexia nervosa. *Am J Obstet Gynecol.* 1973;117:435–449.
9. Halmi KA. Anorexia nervosa: demographic and clinical features in 94 cases. *Psychosom Med.* 1974;36:18–25.
10. Fries H. Secondary amenorrhea, self-induced weight reduction and anorexia nervosa. *Acta Psychiatr Scand.* 1974; 248(suppl):1–70. Munksgaard, Copenhagen: Uppsala University, Sweden. Thesis.
11. Russell GFM, Beardwood CJ. Amenorrhea in the feeding disorders: anorexia nervosa and bulimia. *Psychother Psychosom.* 1970;18:359–364.
12. Russell G. Bulimia nervosa: an ominous variant of anorexia nervosa. *Psychol Med.* 1979;9:429–448.
13. Pyle RL, Mitchell JE, Eckert ED. Bulimia: a report of 34 cases. *J Clin Psychiatr.* 1981;42:60–64.
14. Weiss SR, Ebert MH. Psychological and behavioral characteristics of normal-weight bulimics and normal-weight controls. *Psychosom Med.* 1983;45:293–303.
15. Fairburn CG, Cooper PJ. The clinical features of bulimia nervosa. *Br J Psychol.* 1984;144:238–246.
16. Copeland PM, Herzog DB. Menstrual abnormalities. In: Hudson JI, Pope HG, eds. *The Psychobiology of Bulimia.* Washington, DC: American Psychiatric Press; 1987:31–54.
17. Mello NK. Effects of alcohol abuse on reproductive function in women. *Recent Dev Alcohol.* 1988;6:253–276.
18. Abraham SF, Beumont PJV, Fraser IS, Llewellyn D. Body weight, exercise and menstrual status among ballet dancers in training. *Br J Obstet Gynaecol.* 1982; 89:507–510.
19. Brooks-Gunn J, Warren MP, Hamilton LH. The relation of eating problems and amenorrhea in ballet dancers. *Med Sci Sports Exerc.* 1987;19:41–44.
20. Yahiro J, Glass AR, Fears WB, Ferguson EW, Vigersky RA. Exaggerated gonadotropin responses to luteinizing hormone-releasing hormone in amenorrheic runners. *Am J Obstet Gynecol.* 1987;156:586–591.
21. Shangold MM. Causes, evaluation and management of athletic oligo/amenorrhea. *Med Clin North Am.* 1985;69:83–95.
22. Highet R. Athletic amenorrhea: an update on aetiology, complications and management. *Sports Med.* 1989;7:82–108.
23. Whitaker A, Davies M, Shaffer D, et al. The struggle to be thin: a survey of anorexic and bulimic symptoms in a non-referred adolescent population. *Psychol Med.* 1989;19:143–163.
24. Gritz ER. Cigarette smoking by adolescent females: implications for health and behavior. *Wom Health.* 1984;9:103–115.
25. Boyle MH, Offord DR. Smoking, drinking and use of illicit drugs among adolescents in Ontario: prevalence, patterns of use and sociodemographic correlates. *Can Med Assoc J.* 1986;135:1113–1121.
26. Palmer JH, Ringwalt CL. Prevalence of alcohol and drug use among North Carolina public school students. *J Sch Health.* 1988;58:288–291.
27. Gardner J. Adolescent menstrual characteristics as predictors of gynecological health. *Ann Hum Biol.* 1983;10:31–40.
28. Jones KP, Ravnikar VA, Tulchinsky D, Schiff I. Comparison of bone density in amenorrheic women due to athletics, weight loss, and premature menopause. *Obstet Gynecol.* 1985;66:5–8.
29. Biller BM, Saxe V, Herzog DB, Rosenthal DI, Hozman S, Klibanski A. Mechanisms of osteoporosis in adult and adolescent women with anorexia nervosa. *J Clin Endocrinol Metab.* 1989;68:548–554.
30. Marcus R, Cann C, Madvig P, et al. Menstrual function and bone mass in elite women distance runners: endocrine and metabolic features. *Ann Intern Med.* 1985;102:158–163.
31. Warren MP, Brooks-Gunn J, Hamilton LH, Warren LF, Hamilton WG. Scoliosis and fractures in young ballet dancers: relation to delayed menarche and secondary amenorrhea. *N Engl J Med.* 1986;314: 1348–1353.
32. Frisch RE, Wyshak G, Witschi J, Albright NL, Albright TE, Schiff I. Lower lifetime occurrence of breast cancer and cancers of the reproductive system among former college athletes. *Int J Fertil.* 1987;32:217–225.
33. Winter JSD, Faiman C. The development of cyclic pituitary-gonadal function in adolescent females. *J Clin Endocrinol Metab.* 1973;37:714–718.
34. Apter D. Serum steroids and pituitary hormones in female puberty: a partly longitudinal study. *Clin Endocrinol.* 1980;12:107–120.
35. Metcalf MG, Skidmore DS, Lowry GF, Mackenzie JA. Incidence of ovulation in the years after menarche. *J Endocrinol.* 1982;97:213–219.
36. Brooks-Gunn J, Warren MP, Rosso J, et al. Validity of self-report measures of girls' pubertal status. *Child Dev.* 1987;58:829–841.
37. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* 3rd ed. Washington, DC: American Psychiatric Press;1980:69–71.
38. Garner DM, Olmstead MP, Bohr Y, Garfinkel PE. The Eating Attitudes Test: psychometric features and clinical correlates. *Psychol Med.* 1982;12:871–878.
39. Whitaker A, Johnson J, Shaffer D, et al. Untold troubles: prevalence estimates of selected psychiatric disorders in a secondary school population. *Arch Gen Psychiatry.* 1990;47:487–496.
40. Kandel D. Adolescent marihuana use: role of parents and peers. *Sci.* 1973;181:1067–1069.
41. Bauman KE, Koch GG, Bryan ES. Validity of self-reports of adolescent cigarette smoking. *Int J Addict.* 1982;17:1131–1136.
42. Smart RG, Jarvis GK. Do self-report studies of drug use really give dependable results? *Can J Criminol.* 1981;23:82–92.
43. Single E, Kandel D, Johnson B. The reliability and validity of drug use responses in a large scale longitudinal study. *J Drug Issues.* 1975;5:426–443.
44. Hollingshead AB. The four factor index of social status. New Haven, Conn: Yale University; 1974. Mimeo.
45. Keys A, Fidanza F, Karvonen MJ, Kimura N, Taylor HL. Indices of relative weight and obesity. *J Chronic Dis.* 1972;25:329–343.
46. Killeen J, Vanderburg D, Harlan WR. Application of weight-height ratios and body indices to juvenile populations: the National Health Examination Survey data. *J Chronic Dis.* 1978;31:529–537.
47. Vollman RF. The degree of variability of length of the menstrual cycle in correlation with age of woman. *Gynaecol.* 1956; 142:310–314.
48. Zlatnik FJ, Burmeister LF. Low 'gynecologic age': an obstetric risk factor. *Am J Obstet Gynecol.* 1977;128:183–186.
49. Greenland S. Interpretation and choice of effect measures in epidemiologic analysis. *Am J Epidemiol.* 1987;125:761–768.
50. Kelsey JL, Thompson WD, Evans AS. *Methods in Observational Epidemiology.* New York, NY: Oxford University Press; 1986:39–41, 113–119.
51. Morganstern H. Uses of ecologic analysis in epidemiologic research. *Am J Public Health.* 1982;72:1335–1344.
52. Greenland S. Bias in methods for deriving standardized morbidity ratio and attributable fraction estimates. *Stat Med.* 1984;3:131–141.
53. Walter SD. Calculation of attributable risks from epidemiological data. *Int J Epidemiol.* 1978;7:175–182.
54. Miettinen OS. Stratification by a multivariate confounder score. *Am J Epidemiol.* 1976;104:609–620.
55. Fleiss JL, Williams JBW, Dubro AF. The logistic regression analysis of psychiatric data. *J Psychiatr Res.* 1986;20:145–209.
56. Greenland S. Modeling and variable selection in epidemiologic analysis. *Am J Public Health.* 1989;79:340–349.
57. Pettersson F, Fires H, Nillius SJ. Epidemiology of secondary amenorrhea: I. incidence and prevalence rates. *Am J Obstet Gynecol.* 1973;117:80–86.
58. Venturoli S, Porcu E, Fabbri R, et al. Menstrual irregularities in adolescents: hormonal pattern and ovarian morphology. *Horm Res.* 1986;24:269–278.
59. Johnson C, Lewis C, Love S, Lewis L, Stuckey M. Incidence and correlates of bulimic behavior in female high school population. *J Youth Adolesc.* 1984;13:15–26.
60. Kiriike N, Nishiwaki S, Nagata T, Maeda Y, Kawakita Y. Gonadotropin response to LH-RH in anorexia nervosa and bulimia. *Acta Psychiatr Scand.* 1988;77:420–426.
61. Pirke KM, Dogs M, Fichter MM, Tuschl RJ. Gonadotrophins, oestradiol and progesterone during the menstrual cycle in bulimia nervosa. *Clin Endocrinol.* 1988; 29:265–270.
62. Devlin MJ, Walsh BT, Katz JL, et al. Hypothalamic-pituitary-gonadal function in anorexia nervosa and bulimia. *Psychiatry Res.* 1989;28:11–24.
63. Deuster PA, Kyle SB, Moser PB, Vigersky RA, Singh A, Schoomaker EB. Nutritional intakes and status of highly trained amenorrheic and eumenorrheic women runners. *Fertil Steril.* 1986;46:636–643.
64. Pirke KM, Schweiger U, Laessle R, Dickhaut B, Schweiger M, Waechtler H. Dieting influences the menstrual cycle: vegetarian vs nonvegetarian diet. *Fertil Steril.* 1986;46:1083–1088.
65. Frisch RE. Body fat, puberty and fertility. *Biol Rev.* 1984;59:161–188.
66. Pirke KM, Schweiger U, Lemmel W, Kreig JC, Berger M. The influence of diet-

ing on the menstrual cycle of healthy young women. *J Clin Endocrinol Metab.* 1985;60:1174–1179.
67. Fries H, Nillius SV, Pettersson F. Epidemiology of secondary amenorrhea: II. a retrospective evaluation of etiology with special regard to psychogenic factors and weight loss. *Am J Obstet Gynecol.* 1974; 116:473–479.
68. Weiner H. Psychoendocrinology of anorexia nervosa. *Psychiatr Clin North Am.* 1989;12:187–206.
69. Holmberg NG, Nylander I. Weight loss in secondary amenorrhea. *Acta Obstet Gynecol Scand.* 1971;50:241–246.
70. Knuth JA, Hull MG, Jacobs HS. Amenorrhea and loss of weight. *Br J Obstet Gynaecol.* 1987;84:801–807.
71. Bates GW. Body weight control practice as a cause of infertility. *Clin Obstet Gynecol.* 1985;28:632–644.
72. Kreipe RE, Strauss J, Hodgman CH, Ryan RM. Menstrual cycle abnormalities and subclinical eating disorders: a preliminary report. *Psychosom Med.* 1989;51:81–86.
73. Sloss EM, Frerichs RR. Smoking and menstrual disorders. *Int J Epidemiol.* 1983;12:107–109.
74. Howe G, Westhoff C, Vessey M, Yeates D. Effects of age, cigarette smoking and other factors on fertility: findings in a large prospective study. *Br Med J.* 1985; 290:1697–1700.
75. Wainwright RL. Change in observed birth weight associated with change in maternal cigarette smoking. *Am J Epidemiol.* 1983;117:668–673.
76. Kline J, Stein ZA, Susser M, Warburton D. Smoking: a risk factor for spontaneous abortion. *N Engl J Med.* 1977;297:793–796.
77. Willett W, Stampfer MJ, Bain C, et al. Cigarette smoking, relative weight and menopause. *Am J Epidemiol.* 1983;117:651–658.
78. McKinlay SM, Bifano NL, McKinlay JB. Smoking and age at menopause in women. *Ann Int Med.* 1985;103:350–356.
79. Weisberg E. Invited review: smoking and reproductive health. *Clin Reprod Fertil.* 1985;3:175–186.
80. Wilson E, Emans SJ, Mansfield J, Podolsky C, Grace E. The relationships of calculated percent body fat, sports participation, age, and place of residence on menstrual patterns in healthy adolescent girls at an independent New England high school. *J Adolesc Health Care.* 1984;5: 248–253.
81. Gadpaille WJ, Sanborn CF, Wagner WW. Athletic amenorrhea, major affective disorders and eating disorders. *Am J Psychiatry.* 1987;144:939–942.

Hospital Stays Continue 10-Year Decline

Length of stay continues to decline, reflecting a 10-year trend in shortened hospital stays, according to data in the *1990 Length of Stay By Diagnosis and Operation* series released by Healthcare Knowledge Resources in Ann Arbor, Mich.

Hospital stays show a 17% decline since 1981 when the average length of stay in the United States was 6.7 days. From 1989 to 1990 alone, the national average stay for patients in short-term, general hospitals dropped 1.8%, from 5.6 days to 5.5.

The downward trend in length of hospital stay is even more evident when individual diagnoses and procedures are evaluated. For example, the average length of stay for a normal delivery declined almost 37%, from 3.0 days in 1981 to 1.9 days in 1990. Predelivery education, the advent of LDRs (labor-delivery rooms) and pressure from third-party payers contributed to the shortened stay. Experts predict that the decrease in average length of stay for normal deliveries will stop before reaching 1 day.

Regional differences are also a factor in the average length of stay. Traditionally, hospital stays are highest in the Northeast and North Central regions and lowest in the South and West. This pattern was consistent until 1990, when the average length of stay was lower in the North Central region (5.6 days) than in the South (5.7 days). The average length of stay has remained constant in the South since 1987, whereas the North Central region has shown a steady decline. The lowest average length of stay is in the West at 4.8 days; the highest is in the Northeast at 6.2 days.

Healthcare Knowledge Resources' *Length of Stay* data are drawn from the Commission on Professional and Hospital Activities (CPHA) database and reflects over 4 million discharge records from short-term, general hospitals. Data have been collected since 1963 and are available for separate regions, the whole United States, and also for geriatric and pediatric patients.

For more information on *Length of Stay*, contact Healthcare Knowledge Resources, PO Box 303, Ann Arbor, MI 48106-0303. Within the United States, call toll free 800/521-6210; outside the United States, call 313/930-7830.

Predictors of Smoking Prevalence among New York Latino Youth

Linda Dusenbury, PhD, Jon F. Kerner, PhD, Eli Baker, PhD, Gilbert Botvin, PhD, Susan James-Ortiz, MA, and Ann Zauber, PhD

ABSTRACT

Background. We examined prevalence rates and risk factors for smoking among Latino adolescents, using a multiethnic sample of sixth- and seventh-grade students (n = 3129) in 47 New York City public and parochial schools.

Methods. The students completed questionnaires; self-reported smoking data were collected by means of the "bogus pipeline" technique. The largest group of Latino students (43%) was Puerto Rican; 20% were of Dominican background, 7% were Colombian, and 7% were Ecuadorian. "Current smoking" was defined as smoking at least once per month.

Results. A series of logistic regression analyses indicated that peer influence was the strongest predictor of smoking. Family influence was important as well.

Conclusions. The results are discussed in terms of their implications for prevention. (*Am J Public Health.* 1992;82:55–58)

Introduction

Despite a decline in the prevalence of smoking among adult White males since the Second World War, there has been a steady increase in smoking among adult White females and adult minority group members, and the same patterns have been observed among adolescents.[1] According to the US Office of Minority Health,[2] there are important but poorly understood differences in health behaviors within and among various racial and ethnic groups, and this appears to be particularly true for Latinos.[3] Marcus and Crane[4] conclude that Latino men are as likely to smoke as non-Latino men, and that the rates of smoking-related morbidity and mortality among Latino men are expected to rise in the 1990s. Similarly, Escobedo and Remington[5] report that the rate of smoking among Latino women is increasing.

Determining smoking prevalence among Latino groups has been a problem because different studies have used different samples to estimate prevalence, as well as different operational definitions of "smoking." The only data for Latino adolescents come from the Hispanic Health and Nutrition Examination Survey,[6] which used "current smoking" to explore prevalence. That survey found that approximately 1% of Mexican-American adolescents aged 12 to 14 years smoked, compared with 3.5% for Puerto Ricans. In both groups the rate of smoking was three times higher for girls than for boys.

Austin and Gilbert[7] note in a review article that the correlates of substance use (not including smoking) generally are similar for Latino and non-Latino populations, and include lower socioeconomic status, family instability, peer influence, and family influence. However, little information has been available concerning Puerto Rican youth specifically, and no information has been available concerning Dominicans, Colombians, or Ecuadorians.

In the present study, we examined correlates of adolescent smoking among New York City Latino populations attending public and Catholic schools. Previously reported predictors of smoking in White non-Latino populations were tested with the overall sample of Latino youth available in this study.

Methods

This study was part of an ongoing longitudinal investigation of smoking prevention interventions in New York City schools whose student bodies were more than 25% Latino. The intervention was implemented in the sixth and seventh grades. All sixth- and seventh-graders in English-speaking, mainstream classes in the schools participated in the study. A passive consent procedure was used, and more than 90% of the students completed the baseline survey of smoking behavior. Data for the present study were derived from the baseline survey and were collected prior to the intervention; the sample

Linda Dusenbury, Eli Baker, and Gilbert Botvin, are with Cornell University Medical College, New York, NY. Jon F. Kerner, Susan James-Ortiz, and Ann Zauber are with Memorial Sloan-Kettering Cancer Center, New York, NY.

Requests for reprints should be sent to Linda Dusenbury, PhD, Department of Public Health, Cornell University Medical College, 411 East 69th Street, New York, NY 10021.

This paper was submitted to the journal May 2, 1990, and accepted with revisions July 10, 1991.

TABLE 1—Current Smoking[a] Prevalence among Latino Students, by Gender and Age

Age	Females n	Females %	Females 95% CI	Males n	Males %	Males 95% CI
Puerto Rican						
≤ 12 years	245	5.3	2.5–8.1	251	1.6	0–3.1
13 years	371	8.1	4.9–11.4	253	5.1	2.4–7.9
≥ 14 years	113	20.4	12.9–27.8	145	13.1	7.6–18.6
Other Latino						
≤ 12 years	465	4.1	2.3–5.9	371	1.3	0.2–2.5
13 years	331	3.6	1.6–5.6	309	4.9	2.5–7.3
≥ 14 years	99	6.1	1.3–10.8	127	7.1	2.6–11.6

Note. CI = confidence interval.
[a]"Current smoking" was defined as smoking at least once per month.

consisted of students from both experimental and control schools.

The schools eligible for inclusion in this study were Catholic and public schools in New York City whose student bodies were at least 25% Latino. Forty-seven schools in New York City participated in the intervention study; 36 were Catholic schools and 11 were public schools. The sample was evenly distributed between the two school types, with 2681 students (49%) from Catholic schools and 2820 students (51%) from public schools. Both public and Catholic schools in New York City serve lower-income, urban minority youth.

Subjects

A total of 3129 sixth- and seventh-grade students who identified themselves as Latino or Hispanic on the questionnaire participated in the study. The mean age of the students was 12.61 years (range, from 10 to 18 years). Forty-eight percent of the students were male. Of the total sample, 43% were Puerto Rican, 20% were Dominican, 7% were Colombian, and 7% were Ecuadorian. The remaining 23% of the students identified with another Latino subgroup or with more than one Latino subgroup.

Instrument

The instrument was a questionnaire consisting of 143 items concerning demographics, substance use, and psychosocial factors. It included items on self-reported smoking and the perceived social support for smoking.

Background variables. Twelve items were used to elicit basic background information concerning gender, age, family structure, and academic performance (grades). In addition, students' race was determined by a single item. If students indicated that they were Latino or Hispanic, they were asked to indicate their Latino subgroup status, which included Puerto Rican, Dominican, Colombian, Ecuadorian, and other Latino subgroups and subgroup combinations. Only those students identifying themselves as Latino or Hispanic were included in the present study.

Smoking status. A dichotomous (yes or no) item was used to assess smoking during the last month. A general 11-point scale was used to assess smoking frequency. Data from this 11-point scale were used to group students into three categories: those who had never smoked, those who smoked less than monthly (experimental smokers), and those who smoked monthly to daily (current smokers).

Perceived social influences on smoking. Social influences on smoking were assessed, including (1) smoking by significant others (friends, siblings, and parents), (2) perceived attitudes of parents and friends toward respondents' smoking, and (3) perceived prevalence of general smoking rates among peers and adults (normative expectations). Smoking by friends was measured on a 5-point scale ranging from "none" to "almost all." Smoking by older siblings was measured on a 4-point scale ranging from "none" to "three or more." Smoking by parents was assessed in separate items for the respondent's father and mother ("Does your father [mother] smoke cigarettes?"; responses included "have no father/mother," "no," "used to but quit," and "yes"). The perceived attitudes of parents and friends toward the respondent's smoking were assessed, in separate items, on a 5-point scale ranging from "strongly against it" to "strongly in favor of it." The perceived prevalence of general smoking rates among peers and adults was measured, in two items ("In your opinion, how many people your age [adults] smoke cigarettes?"), on a 6-point scale ranging from "none" to "almost all."

Procedure

The survey was administered by project staff members. The students completed the questionnaire during a regular 40-minute classroom period. Teachers were not involved in the data collection activities and the students were assured that their answers would remain confidential. The students also provided a breath sample for carbon monoxide testing. Although correlations between carbon monoxide levels and self-reports of smoking behavior among students in this age group are typically too low to use as an independent validity check, this "bogus pipeline" procedure has been found to increase the veracity of self-reported smoking data.[9]

Data Analysis

The data were analyzed with SAS software.[10] First, within the major ethnic groups, males and females were compared in terms of smoking prevalence. Next, logistic regressions were used to analyze predictors of smoking status. No significant differences were found between school types; therefore, the two groups were combined.

Results

Table 1 presents the prevalence of current (at least monthly) smoking by ethnic and gender groups. Older students were generally more likely than younger students to be current smokers. In addition, Puerto Rican female students of all ages tended to have higher current smoking prevalence rates than Puerto Rican male students or other female students.

Two logistic regression analyses were carried out to compare adolescents who had never smoked (n = 2180) with adolescents who had smoked at least once (experimental smokers, n = 616), and adolescents who were current smokers (n = 99) with all others (n = 2697). Demographic and prosmoking social-influence variables were entered into the equations simultaneously. Individuals for whom any of the variables in the equation was missing were omitted from the analysis. Latino groups were compared using Puerto Ricans as the reference group. Prosmoking social-influence variables included prevalence of smoking among friends, siblings, and parents; perceived smoking prevalence among youth and

adults generally; and parents' and friends' attitudes about respondents' smoking.

Several demographic variables emerged as significant predictors of experimental smoking (Table 2). Specifically, older students and students with lower academic performance were more likely to have ever smoked.

Latino students who were Dominican, Colombian or Ecuadorian (South American), or from another Latino category were somewhat more likely to have ever smoked than were Puerto Rican students. Thus, even though Puerto Rican girls were more likely to be current smokers (see Table 1), Dominicans, Ecuadorians, and Colombians were more likely to have ever smoked by seventh grade.

Prosmoking social influences were also related to experimental smoking. The general prevalence of smoking among peers was a significant predictor of having smoked: the odds of having smoked were 55% lower for students who reported that few to none of their peers smoked than for students who reported that from more than half to all of their peers smoked.

The major predictor of experimental smoking was the proportion of friends who smoked. Students who reported that from 50% to 100% of their friends smoked were eight times more likely to have ever smoked than were students who reported having no friends who smoked. Number of siblings who smoked and maternal smoking were also predictors. Parents' and friends' attitudes toward respondents' smoking were predictors: less negative attitudes were correlated with greater odds of smoking.

Table 3 presents the significant predictors of current smoking. Although the number of students who met the criterion for current smoking was small (n = 99), the results of this exploratory analysis were largely consistent with the findings of the previous logistic regression analysis. Specifically, age and academic performance were significant predictors of current smoking. In addition, the proportion of friends who smoked was the most important predictor of current smoking, as it was for experimental smoking. Students who reported that between 50% and 100% of their friends smoked were 17 times as likely to be current smokers. Sibling and maternal smoking also increased the odds of current smoking, as they had in the case of experimental smoking, but in the case of current smoking, paternal smoking also emerged as a significant predictor. Parental attitudes again were correlated with current smoking, although friends' attitudes did not emerge as a significant predictor.

TABLE 2—Predictors of Experimental Smoking

Predictor	Odds Ratio	95% Confidence Interval
Age	1.20	1.06–1.35
Gender	.97	.79–1.20
Academic performance		
B's and C's	1.35	1.06–1.73
C's and D's	1.35	1.04–1.76
Latino group (Puerto Rican[a])		
Dominican	1.56	1.17–2.08
South American	1.81	1.32–2.50
Other Latino	1.33	1.01–1.76
Friends smoke (None[a])		
<50%	4.23	3.34–5.37
50% to 100%	8.06	5.96–10.90
Siblings smoke (None[a])		
No siblings	1.16	.90–1.48
One smokes	2.14	1.60–2.87
Two+ smoke	2.48	1.72–3.57
Mother smokes (No[a])		
Quit	1.38	1.00–1.90
Yes	1.36	1.07–1.72
Father smokes (No[a])		
Quit	.80	.58–1.12
Yes	1.14	.91–1.43
General prevalence among peers (>50% to All[a])		
50%	.97	.73–1.29
<50%	.83	.58–1.18
Few or none	.55	.42–.72
Parents' attitude (Strongly negative[a])		
Negative	1.92	1.30–2.83
Neutral or favorable	1.43	1.08–1.90
Friends' attitude (Strongly negative[a])		
Negative	2.32	1.68–3.21
Neutral or favorable	1.81	1.38–2.36

[a]Reference category.

Discussion

This study examined prevalence rates and risk factors for smoking among Latino adolescents, using a multiethnic sample of sixth- and seventh-grade students in New York City. The pattern of smoking among the different Latino subgroups in this study is fairly straightforward, and is consistent with patterns found in previous studies of non-Latino populations. Our findings suggest that age, academic performance, Latino subgroup, and social influences, including friends who smoke as well as family members who smoke, are all related to smoking status. Consistent with Austin and Gilbert's[7] review of predictors of Latino substance use (not including smoking), peer influences were the strongest predictors of smoking for Latinos in the present study.

Several factors may limit the generalizability of the present study. The first concerns the sample. Students may differ from dropouts or nonstudents. Furthermore, the students who participated in this study were those who were enrolled in English-speaking classes. Therefore, this study does not provide information concerning the population of Latino students enrolled in bilingual classes.

The results of this study have several implications for prevention. Predictors of smoking in this Latino sample are also consistent with findings for White non-Latino adolescents[11] and Black non-Latino adolescents.[12] The results suggest that in interventions with sixth- and seventh-graders *in school settings,* the differences between ethnic groups may not be as important as the overall similarities with respect to predictors of smoking. Furthermore, the interventions most likely to prevent smoking would be those that promote peer resistance skills, and the motivation to use those skills, in all groups. This does not mean that cultural differences are not important; the potential reception by the target audience might be enhanced, and the efficacy of

TABLE 3—Predictors of Current Smoking

Predictor	Odds Ratio	95% Confidence Interval
Age	1.30	1.02–1.65
Gender (female[a])	1.05	.66–1.67
Academic performance		
B's and C's	2.65	1.35–5.19
C's and D's	2.32	1.11–4.81
Latino group (Puerto Rican[a])		
Dominican	1.33	.72–2.48
South American	.54	.23–1.28
Other Latino	.66	.34–1.28
Friends smoke (None[a])		
<50%	4.87	2.24–10.60
50% to 100%	16.92	7.84–36.51
Siblings smoke (None[a])		
No siblings	1.40	.74–2.64
One smokes	2.92	1.56–5.49
Two+ smoke	3.42	1.79–6.53
Mother smokes (No[a])		
Quit	.82	.37–1.84
Yes	1.68	1.03–2.73
Father smokes (No[a])		
Quit	1.97	.95–4.08
Yes	1.70	1.02–2.83
General prevalence among peers (>50% to All[a])		
50%	1.03	.59–1.78
<50%	.60	.26–1.39
Few or none	.53	.28–1.00
Parents' attitude (Strongly negative[a])		
Negative	1.95	.90–4.22
Neutral or favorable	1.77	1.05–2.99
Friends' attitude (Strongly negative[a])		
Negative	1.04	.44–2.46
Neutral or favorable	1.80	.93–3.49

[a]Reference category.

the intervention maximized, by material sensitive to a whole range of cultural orientations. Further study is needed to explore this issue. The results of this study also suggest that smoking prevention in schools with large Latino populations might be enhanced by programs that also address family influences on student smoking.

This study extends previous knowledge of the prevalence and concurrent predictors of smoking among Latino youth. Additional research will be necessary to develop a complete model for understanding the etiology of smoking in these groups. For example, the relationship of language and acculturation to smoking is complex and requires a comprehensive examination. Other cultural factors that may play a role in the etiology of smoking in Latino youth and require additional research include community influences, microcultures, and cultural beliefs and attitudes concerning health.[13] Finally, individual psychosocial factors, such as knowledge and attitudes, were not included in this study, but these may also emerge as powerful predictors of smoking in Latino youth. □

Acknowledgments

This research was supported by Grant 1 R18 CA 39280 from the National Cancer Institute. For their support and cooperation in this project, we would like to thank the administrations and the schools of the Archdiocese of New York, the Diocese of Brooklyn and Queens, and the New York City Board of Education.

References

1. U.S. Department of Health and Human Services. *Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General.* U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. DHHS Publication No. (CDC) 89-8411, 1989.
2. Gold RS, Thomas SB, Davis D, Overby L, Kelly M. *The Literature on Prevention in Minority Communities: Some Lessons to be Learned.* US Office of Minority Health, Minority Health Research Laboratory; 1987. Unpublished manuscript.
3. Escobedo LG, Remington PL, Anda RF: Long-term secular trends in initiation of cigarette smoking among Hispanics in the United States. *Public Health Rep.* 1989;104:583–587.
4. Marcus AC, Crane LA. Smoking behavior among U.S. Latinos: an emerging challenge for public health. *Am J Public Health.* 1985;75:169–172.
5. Escobedo LG, Remington PL. Birth cohort analysis of prevalence of cigarette smoking among Hispanics in the United States. *JAMA.* 1989;26:66–69.
6. Haynes SG, Harvey CR. *Data from Hispanic HANES.* Presented at the National Cancer Institute grantee workshop on smoking prevention and cessation among U.S. Hispanic populations; January 21, 1988; Rockville, Md.
7. Austin GA, Gilbert MJ. Substance abuse among Latino youth. *Prev Res Update.* Spring 1989;3:1–28.
8. Botvin GJ, Dusenbury L. Substance abuse prevention and the promotion of competence. In: Bond L, ed. *School-Based Prevention.* Hanover, NH: University Press of New England; 1989. Vermont Conference of Primary Prevention of Psychopathology Series.
9. Evans RI, Hansen WB, Mittelmark MD. Increasing the validity of self-reports of behavior in a smoking in children investigation. *J Applied Psych.* 1979;62:521–523.
10. *SAS/STAT Software: CALIS and Logistic Procedures.* Release 6.04. Cary, NC: SAS Institute, Inc; 1990. SAS Technical Report P.200.
11. Botvin GJ, McAlister A. Cigarette smoking among children and adolescents: causes and prevention. In: Arnold CB, ed. *Annual Review of Disease Prevention.* New York, NY: Springer; 1982:222–249.
12. Botvin GJ, Batson H, Witts-Vitale S, Bess V, Baker E, Dusenbury L. A psychosocial approach to smoking prevention for urban Black youth. *Public Health Rep.* 1989;104: 573–582.
13. Ramirez AG, MacKellar DA, Gallion K. Reaching minority audiences: a major challenge in cancer reduction. *Cancer Bulletin.* 1988;40:334–343.

Variations in Asthma Hospitalizations and Deaths in New York City

Willine Carr, DrPH, Lisa Zeitel, MPH, and Kevin Weiss, MD

ABSTRACT

Background. Recent reports have identified New York City as having asthma mortality rates that are substantially higher than expected based on US rates. This study investigates the problems of asthma morbidity and mortality in New York City.

Methods. Data on asthma hospitalizations (1982 to 1986) and deaths (1982 to 1987) among persons aged 0 to 34 years were studied. Descriptive and multivariate techniques were used to examine differences in rates among subgroups and across geographic areas.

Results. The average annual hospitalization rate was 39.2 per 10 000; the mortality rate was 1.2 per 100 000. Hospitalization and death rates among Blacks and Hispanics were 3 to 5.5 times those of Whites. Large geographic variations in hospitalizations and mortality occurred. Asthma hospitalization and mortality rates were highly correlated ($r = .67$), with the highest rates concentrated in the city's poorest neighborhoods. Household income, percentage of population Black, and percentage of population Hispanic were significant predictors of area hospitalization rates (adjusted $R^2 = .75$).

Conclusion. These findings provide a basis for focusing investigations of the causes of variations in asthma outcomes and targeting interventions to reduce the disproportionate morbidity and mortality borne by poor and minority populations. (*Am J Public Health.* 1992;82:59–65)

Introduction

Between 9 and 12 million persons in the United States have asthma, a chronic disease that is a major cause of disability for many persons and a cause of death for some.[1] Asthma is often associated with long-term psychological and financial burdens and is a leading reason for the utilization of health services.[2]

Although the past 20 years have witnessed major advances in the treatment and control of asthma, hospitalizations for asthma are frequent and have increased substantially; asthma mortality has also increased. Evans and colleagues found that between 1965 and 1983, hospitalizations of adults for asthma increased nationally by approximately 50%.[1] Gergen and Weiss reported that asthma hospitalizations of children increased by 47% from 1979 to 1987.[3] From 1980 to 1987, U.S. asthma mortality rates among all ages increased by 31%.[4]

Recent evidence suggests that these increasing trends in asthma hospitalizations and mortality are not occurring evenly across the United States.[5] The trends are to some extent driven by the rates in a few large cities. Significant among these is New York City, where, with less than 3% of the nation's population, approximately 6% of all asthma hospitalizations and 7% of all asthma deaths occurred in 1986. From 1982 to 1985, the number of asthma deaths in New York City among persons aged 5 to 34 was more than three times the number that would be expected based on national rates.[6] Other studies have shown large geographic disparities in asthma hospitalization rates among counties,[7] cities,[8] and hospital market areas.[9] This study is the first to characterize variations in both asthma hospitalizations and asthma deaths among racial/ethnic subgroups and across geographic areas in a single city, and to examine the extent to which area-level sociodemographic factors account for geographic differences.

Methods

Study Population

Asthma hospitalizations and deaths among New York City residents aged 0 through 34 were studied. Hospitalizations were examined for persons with a principal diagnosis of asthma (ICD-9 CM codes 493–493.91) who were discharged from nonfederal general care hospitals in the city during the years 1982 through 1986. Deaths were studied for persons with asthma (493–493.91) listed as the underlying cause of death during the years 1982 through 1987. The additional year for mortality data was included to increase the number of events.

Data Sources

Computerized data on all hospital discharges (asthma and nonasthma) were obtained from the New York State Department of Health. Each discharge record included the patient's age, race/ethnicity, zip code of residence, type of admission, diagnoses and procedures,

Willine Carr and Lisa Zeitel are with the United Hospital Fund of New York. Kevin Weiss is with the Departments of Health Care Sciences and Medicine, George Washington University Medical Center, Washington, DC.

Requests for reprints should be sent to Willine Carr, DrPH, Office of Intergovernmental Relations, Health and Hospitals Corporation, 125 Worth Street, Room 505, New York, NY 10013.

This paper was submitted to the journal November 2, 1990, and accepted with revisions May 2, 1991.

TABLE 1—Hospitalizations for Asthma among New York City Residents Aged 0–34 Years (1982 to 1986)

	Number of Hospitalizations	Average Annual Rate[a]	95% Confidence Interval	Percent Change in Rate 1982–1986
All New York City residents aged 0–34	74 292	39.2	38.7–41.7	17.4
Age				
Under 5 years	23 855	101.5	93.6–109.3	24.4
5–9 years	12 348	55.3	51.7–58.8	8.9
10–19 years	16 856	31.5	30.9–32.1	1.0
20–34 years	21 233	23.5	20.7–26.4	29.0
Sex				
Male	37 629	40.6	37.7–43.6	21.8
Female	36 663	37.8	35.6–40.0	13.0
Race/Ethnicity				
White, non-Hispanic	9 887	12.2	11.7–12.7	5.1
Black, non-Hispanic	31 101	59.9	55.3–64.5	21.1
Hispanic	29 633	62.9	59.1–66.6	8.0

Note. Data for racial/ethnic groups do not add to total because persons categorized as "other" are not shown in the racial/ethnic details.
[a]Per 10 000 population.

length of hospital stay, and expected payment source. The data represent hospitalization events and not individual persons. It was not possible to identify multiple admissions of the same patient.

Mortality information came from vital certificates data compiled by the New York City Department of Health and included the decedent's age, race/ethnicity, and zip code of residence.

Information on population characteristics came from the 1980 U.S. Census. Rates were calculated using the U.S. Census STF-3B data file. Sociodemographic measures used in the multivariate analyses were obtained from a census file prepared by the Claritas Corporation.[10]

Data Analysis

Using descriptive and multivariate techniques, this study focused on variations in asthma hospitalizations and deaths among subgroups and across geographic areas of New York City.

Subgroup analysis. Data were examined for both sexes, for four age groups under age 35, and for racial/ethnic groups categorized as (1) White, non-Hispanic; (2) Black, non-Hispanic; (3) Hispanic; and (4) other.

Rates of hospitalization and mortality were studied at the citywide level for each sex, age, and racial/ethnic group. Since hospitalization rates were essentially unaltered by age and sex adjustment, unadjusted rates are presented. Ninety-five percent confidence intervals for hospitalization and mortality rates were calculated.[11,12] Chi-square tests were used to test the significance of the differences in changes in hospitalization rates among racial/ethnic groups over time.[11] The significance of differences in length of stay among discharges with different payers was tested using analysis of variance.[13]

Small-area analysis. Hospitalizations and deaths were also studied for populations residing in different geographic areas, including the city's 5 boroughs or counties, its 41 neighborhoods, and its 165 residential zip code areas. The populations of persons aged 0 to 34 ranged from 10 090 to 259 515 in the neighborhoods and from 1629 to 67 411 in the zip code areas. Unadjusted rates of hospitalization and death, with ninety-five percent confidence intervals (CIs), were calculated for each geographic area. (Age and sex adjustments altered the rates very little.) At the neighborhood level, the Pearson's product moment correlation coefficient was used to establish the association between neighborhood-level asthma hospitalization and death rates and the association between these outcomes and the income and racial/ethnic compositions of neighborhoods.

Multiple linear regression models were developed, using the SPSS PC Plus Regression procedure, to define the extent to which poverty and race/ethnicity explained geographic variations in hospitalization rates.[14] Zip code areas (N = 165) were used to maximize statistical power and the homogeneity of populations within the geographic areas. The dependent variables were hospitalization rates for persons aged 0 to 34 and for four age-specific groups.

The independent variables used in the regression models were zip code–level measures of socioeconomic status, including median household income, percentage of the population that is Black, and percentage of the population that is Hispanic. Different models were developed by specifying the order in which the explanatory variables were entered. The model shown was selected because it achieved the highest predictive power. Transformations were applied to normalize independent and dependent variables that had skewed distributions; the transformation selected in each case was the one that minimized skewness. Although each pair of predictors was significantly correlated the relationships among the three predictors were not sufficiently high to create a multicollinearity problem in the regression analysis. Residuals in each model were randomly distributed, and thus were not correlated with any of the independent variables.

Results

Asthma Hospitalizations

From 1982 through 1986, New York City residents below age 35 had an average of 14 856 asthma hospitalizations per year and an average annual hospitalization rate of 39.2 per 10 000 population (Table 1). The highest hospitalization rate, 101.5 per 10 000 population, occurred among children under 5 years of age; rates declined to a low of 23.5 per 10 000 for persons aged 20 to 34. Nearly 90% of asthma hospitalizations were unscheduled, primarily resulting from admissions through hospital emergency rooms. Only 0.1% of asthma hospitalizations ended in death.

Blacks and Hispanics accounted for 81.8% of all asthma hospitalizations. The average annual rate per 10 000 population was 12.2 among Whites, 59.9 among Blacks, and 62.9 among Hispanics; Blacks and Hispanics, therefore, had rates five times as high as those of Whites. For each of the age-specific cohorts, minorities were found to have rates at least three times as high as those of Whites.

Asthma hospitalizations increased from 35.8 per 10 000 population in 1982 to 43.1 per 10 000 in 1986, a growth of 17.4% (Table 1). This increase occurred at a time when total hospitalizations for persons under age 35 were declining. In 1982, hospitalizations for asthma constituted 4.4% of all hospitalizations of persons under 35; by 1986, asthma hospitalizations constituted

6.3% of all hospitalizations in this age group. (Total hospitalizations exclude routine obstetric and newborn hospitalizations.) The largest increases in asthma hospitalization rates occurred among persons under age 5 (24.4%) and among persons aged 20 through 34 (29.0%). The increase in the rate among males was 1.7 times the increase for females. All three racial/ethnic groups experienced increases; however, the increase was significantly larger for Blacks than for Hispanics ($\chi^2 = 19.3$, $df = 1$) and Whites ($\chi^2 = 15.2$, $df = 1$). One of the largest race-specific increases (46.4%) occurred among White children under age 5, although the absolute value of this rate in 1986 was still substantially below that of Black and Hispanic children (Figure 1). Black individuals aged 20 to 34 also experienced a very large increase (49.3%) in this rate.

Asthma accounted for substantial use of hospital and third-party-payer resources. Asthma hospitalizations of persons under age 35 amounted to 302 734 days of patient care, 4.1 days per discharge. Length of stay was significantly longer for hospitalizations with Medicaid as the expected source of payment than for hospitalizations with other expected sources of payment ($F = 50.4$, $df = 1$, $p = .0001$). Medicaid was the expected source of reimbursement for 63% of asthma patient days of care, versus 51% of all patient days. (Total patient days exclude routine obstetric and newborn days.)

Asthma Mortality

From 1982 to 1987, a total of 271 deaths due to asthma occurred among New York City residents aged 0 to 34; this represents an average annual rate of 1.2 deaths per 100 000 population (Table 2). About 30% of all asthma deaths occurred among hospitalized asthma patients.

The majority of asthma deaths (68.3%) were among persons aged 20 to 34. As with asthma hospitalizations, more than three fourths (76.2%) of asthma deaths occurred among Blacks and Hispanics. Deaths among Blacks occurred at an average annual rate of 2.2 per 100 000 population, which was 5.5 times the asthma death rate among Whites; the asthma mortality rate among Hispanics, 1.3 per 100 000, was 3 times the White rate.

Asthma mortality peaked at 63 deaths in 1985 (1.7 per 100 000 population, CI = 1.3–2.1), up from 35 in 1982 (0.9 per 100 000, CI = 0.7–1.3). After 1985, mortality rates declined overall and for each of the four age and three racial/ethnic groups.

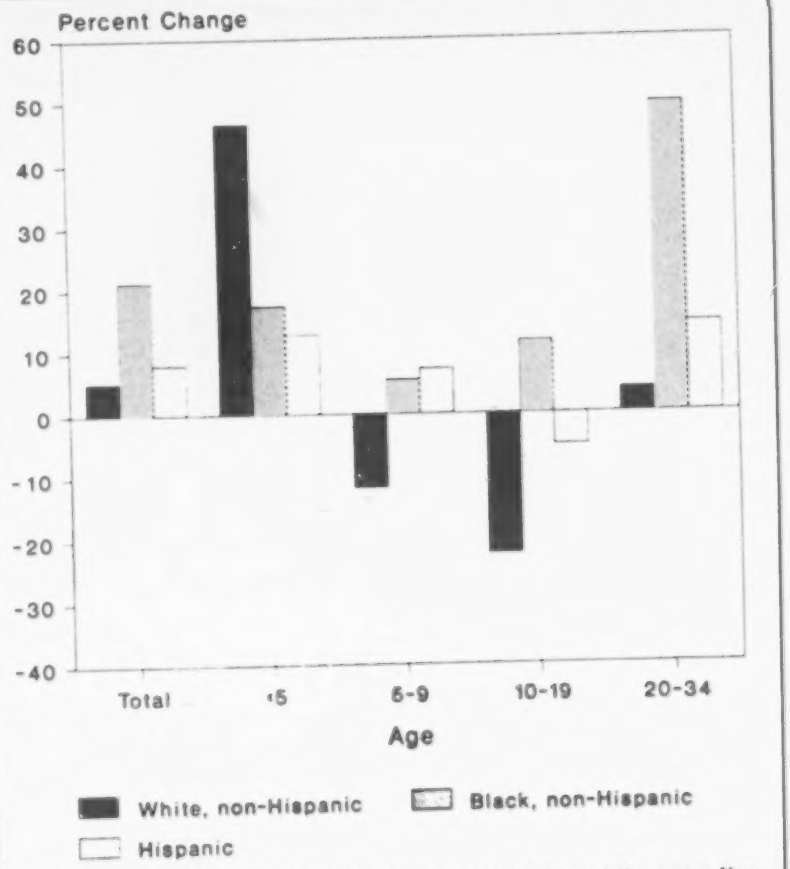


FIGURE 1—Change in asthma hospitalization rates per 10 000 population among New York City residents aged 0–34 years, by age and race/ethnicity, 1982–1986.

TABLE 2—Deaths from Asthma among New York City Residents Aged 0–34 Years 1982–1987

	Number of Deaths	Average Annual Rate[a]	95% Confidence Interval
All New York City residents aged 0–34	271	1.2	1.1–1.3
Age			
Under 5 years	15	0.5	0.3–0.9
5–9 years	9	0.3	0.2–0.7
10–19 years	62	1.0	0.8–1.3
20–34 years	185	1.7	1.5–2.0
Race/Ethnicity			
White, non-Hispanic	42	0.4	0.3–0.6
Black, non-Hispanic	135	2.2	1.8–2.6
Hispanic	73	1.3	1.0–1.6

Note. Data for racial/ethnic groups do not add to total because persons categorized as "other" are not shown in the racial/ethnic details.
[a]Per 100 000 population.

Geographic Variations

Neither asthma hospitalizations nor deaths were distributed evenly across the city. Among New York's five boroughs, average annual asthma hospitalization rates ranged from 65.8 per 10 000 population in the Bronx—1.7 times the citywide rate—to 21.3 per 10 000 in Queens—about

TABLE 3—Average Annual Asthma Hospitalization and Death Rates among New York City Residents Aged 0–34 Years, by Borough

	Hospitalizations (1982–1986)		Deaths (1982–1987)	
	Rate per 10 000 Population	95% Confidence Interval	Rate per 100 000 Population	95% Confidence Interval
All New York City residents aged 0–34	39.2	36.7–41.7	1.2	1.1–1.3
Bronx	65.8	61.3–70.3	1.7	1.3–2.2
Brooklyn	41.6	36.2–47.0	1.4	1.1–1.7
Manhattan	38.8	36.0–41.6	1.2	0.9–1.6
Queens	21.3	18.9–23.6	0.6	0.4–0.9
Staten Island	22.8	19.2–26.5	0.2	0.1–0.8

FIGURE 2—Average annual asthma hospitalization rates per 10 000 population among New York City residents aged 0–34 years, by neighborhood, 1982–1986.

TABLE 4—Zero Order Correlation Matrix for Independent Variables Used in Regression Analyses (Persons Aged 0–34)

	Median Household Income	Percent Population Black[a]	Percent Population Hispanic[a]
Median household income	1.00		
Percent population Black	−0.54	1.00	
Percent population Hispanic	−0.69	0.41	1.00

[a]Log transformation was used to normalize distribution.

one half the citywide rate (Table 3). The average annual asthma mortality rates were higher in the Bronx (1.7 per 100 000 population) and in Brooklyn (1.4 per 100 000) than in other boroughs.

Differences in asthma hospitalization rates among the city's neighborhoods are striking (Figure 2). The average annual hospitalization rate in the neighborhood with the highest rate was nearly 16 times the rate in the neighborhood with the lowest rate. The highest rate, 115.0 per 10 000 population (CI = 105.7–124.4), occurred in East Harlem in Manhattan, where 93% of the population is Black or Hispanic and where the 1979 median household income was under $11 000. The lowest rate, 7.2 per 10 000 (CI = 5.8–8.7), occurred in Greenwich Village–Soho in Manhattan, a neighborhood that is 70% white with a 1979 median household income that exceeded $19 000. Neighborhood asthma hospitalization rates showed a high negative correlation with neighborhood median household income ($r = -.80$), and were positively correlated with percentage of the population Black ($r = .74$) and Hispanic ($r = .75$).

Asthma death rates also varied greatly across areas (Figure 3). Neighborhood asthma death rates, like hospitalizations, were negatively correlated with median household income ($r = -.65$), and positively correlated with percentage of population Black ($r = .56$) and percentage Hispanic ($r = .59$). The highest asthma mortality rate also occurred in East Harlem, where the average annual death rate was 4.9 per 100 000 population (CI = 3.0–7.9), more than four times the citywide rate. Of the asthma deaths for which the neighborhood of the deceased could be identified (N = 252), 45% occurred in only 6 of the city's 41 neighborhoods, neighborhoods that accounted for only 21% of the city's population. The 1979 median household income in these neighborhoods was under $13 000.

Variations in asthma hospitalizations and deaths follow virtually the same geographic pattern. High rates of both asthma hospitalization and asthma death are strongly associated with both race/ethnicity and poverty. Not surprisingly, then, asthma hospitalization rates by neighborhood are highly correlated with asthma mortality rates ($r = .67$).

Multiple Regression Analysis

Three independent variables were examined for their contribution to asthma hospitalization rates. (See Table 4 for correlation matrix.) Among persons aged 0 to

TABLE 5—Results of Regression Analyses for Asthma Hospitalizations

Age Group	Variable[a]	Change in R^2	Zero Order Correlation Coefficient	Partial Correlation Coefficient	β	*P*
0–34 years (Adjusted R^2 = .75)	Median household income	.57	−.75	−.47	−.39	<.0001
	% Population Black	.18	.76	.64	.49	<.0001
	% Population Hispanic	.01	.61	.19	.14	<.05
0–4 years (Adjusted R^2 = .68)	Median household income	.48	−.69	−.38	−.34	<.0001
	% Population Black	.20	.75	.62	.53	<.0001
	% Population Hispanic	.00	.54	.12	.09	NS[b]
5–17 years (Adjusted R^2 = .63)	Median household income	.49	−.70	−.34	−.33	<.0001
	% Population Black	.12	.67	.48	.39	<.0001
	% Population Hispanic	.03	.62	.28	.24	<.001
18–34 years (Adjusted R^2 = .67)	Median household income	.51	−.71	−.41	−.38	<.0001
	% Population Black	.16	.72	.56	.47	<.0001
	% Population Hispanic	.01	.57	.14	.11	NS[b]

Note. The change in R^2 in column 3 reflects the effect of adding a variable to the model. The remaining statistics and probability values are for the final model when all three variables have been entered.
[a]% Population Black and % Population Hispanic were log transformed.
[b]NS = not significant (>.05).

34, nearly three quarters of the geographic variation in asthma hospitalization rates among zip code areas can be accounted for by two of the three ecologic predictors examined: median household income and percentage of the population that is Black (see Table 5). Median household income is significantly negatively related to an area's asthma hospitalization rate ($\beta = -.39$, $P < .0001$), and the percentage of the population that is Black has a significant positive effect on the hospitalization rate ($\beta = .49$, $P < .0001$). The addition of the Hispanic variable to the model, although significant ($\beta = .14$, $P < .05$), increased the amount of variance accounted for by less than 1%. No significant interaction effect due to income was found when interaction terms were included in the model.

Although median household income and percentage Black are both needed to maximize the model's predictive power, each variable on its own is an important determinant of the asthma hospitalization rate in an area. When either of these variables was included as the sole independent variable, it explained about 57% of the disparity in hospitalization rates.

The age-specific models explain between 63% and 68% of the total variation in asthma hospitalization rates. For asthma hospitalization rates of individuals aged 0 to 4 and 18 to 34, the significant predictors were income and percentage Black. However, for the age group 5 to 17, percentage Hispanic also made a significant contribution to explaining the geographic variation.

Discussion

This study clearly demonstrates that the adverse asthma outcomes hospitalization and death are a large health problem in New York City, particularly for Blacks and Hispanics and for residents of low-income areas. These results are consistent, in part, with findings of previous research. Wennberg et al. identified asthma as a discharge diagnosis with a high degree of variation across hospital market areas

FIGURE 3—Average annual asthma death rates per 100 000 population among New York City residents aged 0–34 years, by neighborhood, 1982–1987.

in Maine.[9] In a Maryland study, Wissow and colleagues found large differences in childhood asthma hospitalizations across the state. Their study demonstrated that the greater risk of hospitalization among Black children could be explained largely by poverty rather than race.[7] Perrin et al. showed marked differences in childhood asthma hospitalizations among three cities—Boston, New Haven, and Rochester—and also hypothesized that these differences were related to differences in socioeconomic status.[8]

The present study explicates the correlates and predictors of two asthma outcomes among children and young adults within a single city. The results demonstrate both the high correlation of asthma hospitalization and mortality rates and the strong association of both of these outcomes with community poverty. Although poverty probably accounts for much of the association between asthma hospitalizations and deaths, other possible reasons for this association need to be examined. In addition, the findings indicate that, at the area level, race/ethnicity plays a definite role in asthma hospitalizations when income level is controlled, suggesting that the higher rates of hospitalization found in minority communities are not entirely a function of community poverty. Similarly, the percentage of the population that is Hispanic is significantly associated with asthma hospitalizations in one age group (5- to 17-year-olds), even when the effects of both income and percentage Black are controlled. This unique significance of race/ethnicity should be taken into account in planning epidemiologic studies and health programs for the control of asthma. The large increase in the hospitalization rate for White children under age 5 requires further investigation, but it may be due, in part, to changes in physician recognition and understanding of asthma in this age group.

The high degree of geographic and subgroup variation in asthma hospitalizations and deaths may be driven by differences in the underlying prevalence or case severity for this disease or by differences in environmental exposures. Evans et al. found that asthma is slightly more common in Black than in White populations, especially in younger age groups.[1] The prevalence of asthma among Hispanics is not well documented, although Hispanic children of Puerto Rican origin may have a higher prevalence than non-Hispanic Whites or Blacks.[15] Researchers have found no or only very slight differences in the prevalence of asthma between those above and those below the poverty level.[1] With respect to case severity, there is no information on trends in asthma severity among segments of the U.S. population or on whether asthma is a more severe disease among minorities and poor persons than among others. Finally, the extent and impact of differential exposures among subgroups and across geographic areas to the indoor and outdoor pollutants associated with asthma morbidity are not well established.[16–18]

Even if significant differences in disease prevalence, case severity, and environmental exposures associated with poverty or race/ethnicity are hypothesized, they are unlikely to account fully for the very large disparities in asthma hospitalizations and deaths observed in this study. Clearly, other potential explanations for these differences must be examined, including the role of the health care system.

Asthma is a highly manageable disease. Yet, it has been suggested that inadequate medical management of asthma is common; management tends to be poorest and morbidity, and possibly mortality, highest among patients with the poorest understanding of the disease and its treatment.[19] Given the reversibility of asthma symptoms and the availability of apparently effective treatment strategies,[20] asthma hospitalization and deaths are potentially avoidable outcomes. An excess of these negative outcomes among minorities and the poor may signal possible problems in the structure and functioning of the health care delivery system.[21]

Differences in the supply of health care resources have been hypothesized to account for geographic variations in health services utilization and health outcomes, although, to date, researchers have not documented this relationship with asthma hospitalizations. Wissow et al., for example, found that variations in childhood asthma discharge rates in Maryland were not associated with the supply of hospital beds or with the ratio of population to primary-care physicians.[7] The effect of supply factors was not examined in the present study. However, given the vast supply and distribution of health care resources in New York City, it is reasonable to hypothesize that rather than resource supply per se, the types, configuration, and organization of providers and services available to different population groups may have an impact on asthma outcomes. For example, German and colleagues found that different types of usual sources of health care were related to whether inner-city children in Baltimore received regular care for asthma.[22] The populations found to be at high risk for asthma hospitalization and death in the present study are heavily dependent on hospital outpatient departments and emergency rooms for their ambulatory care.[23]

Differences in medical practice styles, reflecting the exercise of physicians' discretion in the way illnesses are treated, are also important determinants of geographic variation in hospital utilization for many medical conditions.[9] Whether practice styles significantly influence within-city variations in admissions for asthma is an important question in light of the fact that asthma is the leading reason for hospitalization among the age group studied (excluding routine obstetric and newborn hospitalizations). This question is especially important for cities that have overburdened hospital systems with excessively high occupancy rates (such as New York),[24] and should be of particular concern to Medicaid policymakers, since Medicaid is by far the major payer for asthma inpatient services for the age group studied. The observed correlation of asthma hospitalizations and asthma deaths may indicate, however, that hospitalizations are a reasonable indicator of asthma morbidity, and that differences in rates of hospitalization are not primarily a function of differences in admission practices.

The data used in this study have certain limitations. Errors and inconsistencies in the coding of principal diagnoses and underlying causes of death may exist, although accurate identification of asthma morbidity and mortality is least problematic among the age group studied.[25,26]

Another limitation is that identification and coding of race and ethnicity are subject to error and the coding differs for the two data sets used. The mortality data separately identify race and ethnicity, whereas the hospitalization data do not. This difference may account for some of the ethnic patterns shown, such as the higher hospitalization rate but lower mortality rate of Hispanics than of Blacks. Further, a lack of accurate population counts by race/ethnicity at the zip code level precluded direct examination of racial/ethnic patterns of asthma outcomes within small geographic areas.

The hospitalization data represent episodes of care and are not person-specific. It was not possible to determine whether observed differences in rates among racial/ethnic and geographic sub-

groups were due to higher numbers of individuals being hospitalized or a higher rate of readmission among some subgroups. This problem is being investigated by the authors, using a special data set that links individuals' hospitalizations over time.

Finally, the small-area data cannot be used to infer individual-level relationships because of the risk of ecological fallacies.[27] However, as suggested by Kerner and colleagues and as demonstrated in this study, small-area data have valuable uses in identifying high-risk areas and subgroups toward which interventions should be targeted.[28] The intracity findings reported here provide a basis for more detailed and narrowly focused investigations of the factors that contribute most significantly to determining the outcomes for minorities and residents of low-income areas of a common and—usually—relatively benign disease. Such investigations should lead to the design and targeted implementation of public health policies and comprehensive clinical interventions that can reduce the disproportionate burden of asthma morbidity and mortality borne by minorities and the poor. □

Acknowledgments

This study was supported in part by a Public Health Service Epidemiology Training Grant (Dr Weiss).

We acknowledge with gratitude Howard Andrews, PhD, of Columbia University for his statistical consultation and helpful comments on the manuscript. We wish to thank Richard Kaslow, MD, of the National Institute of Allergy and Infectious Diseases, NIH, and Bruce C. Vladeck, PhD, of the United Hospital Fund of New York for their reviews of the manuscript and useful suggestions. We also thank Emily Goodwin, David Gould, PhD, Phyllis Brooks, Toni Heisler, and Melvin I. Krasner, PhD, of the United Hospital Fund for their assistance and advice. Also acknowledged are Jacob Feldman, PhD, and Patricia Knapp of the National Center for Health Statistics for their support in the conduct of the study. We acknowledge the assistance of the Bureau of Biostatistics of the New York City Department of Health in providing mortality data. The skillful secretarial assistance of Liza Buffaloe and Patricia Glover is also greatly appreciated.

References

1. Evans R, Mullally DI, Wilson RW, et al. National trends in morbidity and mortality of asthma in the U.S. *Chest.* 655–754. (Suppl.)
2. Weiss KB, Zimmerman EM, Gergen PJ, Hodgson T. *Special Report: The Socioeconomic Impact of Immunologic and Allergic Disease in the United States.* Report of the Task Force on Immunology and Allergy. Washington, DC: National Institute of Allergy and Infectious Disease, NIH; in press. DHHS publication.
3. Gergen PH, Weiss KB. Changing patterns of asthma hospitalization among children: 1979 to 1987. *JAMA.* 1990;264:1688–1692.
4. Asthma—United States, 1980–1987. *MMWR.* 1990;39:493–497.
5. Weiss KB, Wagener DK. Geographic variation in US asthma mortality: small-area analyses of excess mortality 1981–1985. *Am J Epidemiol.* 1990;132:S107–S115. (Suppl.)
6. Weiss KB, Wagener DK. Changing patterns of asthma mortality: identifying target populations at risk. *JAMA.* 1990;264:1683–1687.
7. Wissow L, Gittelshon AM, Szklo M, Starfield B, Mussman M. Poverty, race, and hospitalization for childhood asthma. *Am J Public Health.* 1988;78:777–782.
8. Perrin JM, Homer CJ, Berwick DM, Woolf AD, Freeman JL, Wennberg JE. Variations in rates of hospitalization of children in three urban communities. *N Engl J Med.* 1989;320:1183–1187.
9. Wennberg JE, McPherson L, Caper P. Will payment based on diagnosis-related groups control hospital costs? *N Engl J Med.* 1984;311:295–300.
10. Claritas Partners LP. *REZIDE File.* Alexandria, Va: Claritas Corporation; 1980.
11. Fleiss JL. *Statistical Methods for Rates and Proportions.* 2nd ed. New York, NY: John Wiley & Sons Inc; 1981:14, 19.
12. Hayes WL. *Statistics.* 4th ed. New York, NY: Holt, Rinehart & Winston; 1988:235–239.
13. Fleiss JL. *Statistical Methods for Rates and Proportions.* 2nd ed. New York, NY: John Wiley & Sons Inc; 1981:225.
14. Neter J, Wasserman W, Kutner M. *Applied Linear Regression Models.* Homewood, Ill: Richard E. Irwin; 1983.
15. Carter O, Gergen P, Lecca P. Reported asthma among Puerto Rican, Mexican-American and Cuban-American children: HHANES, 1982–1984. Presented at the 115th Annual Meeting of the American Public Health Association; October 19, 1987; New Orleans.
16. Sporik R, Holgate ST, Platts-Mills TA, Cogswell JJ. Exposure to house-dust mite allergen (Der p I) and the development of asthma in childhood—A prospective study. *N Engl J Med.* 1990;323:502–507.
17. Dockery DW, Speizer FE, Stram DO, Ware JH, Sengler JD, Ferris GH. Effects of inhalable particles on respiratory health of children. *Am Rev Respir Dis.* 1989;39:587–594.
18. Goldstein IF, Reed CE, Swanson MC, Jacobson JS. Aeroallergens in New York inner-city apartments of asthmatics. *Adv in Aerobiol.* 1987;51:133–137.
19. Sly RM. Effects of treatment on mortality from asthma. *Ann Allergy.* 1986;56:207–212.
20. McFadden ER. Asthma. In: Braunwald E, ed. *Harrison's Principles of Internal Medicine.* 11th ed. New York, NY: McGraw Hill; 1987:1060–1065.
21. Rutstein D, Berenberg W, Chalmers TC, Child CG, Fishman AP, Perrin EB. Measuring the quality of medical care: a clinical method. *N Engl J Med.* 1976;294:582–588.
22. German PS, Skinner EA, Shapiro S, Salkever DS. Preventive and episodic health care of inner-city children. *Community Health.* 1976;2:92–106.
23. Carr W, Schop JA, Cohen S. *New York City Health Facts: A Profile from the National Health Interview Survey.* New York, NY: United Hospital Fund of New York; 1989.
24. Myers LP, Spitz B, Vladeck B, et al. *New York City's Health Occupancy Crisis: Caring for a Changing Patient Population.* New York, NY: Bigel Institute for Health Policy and United Hospital Fund of New York; August 1988.
25. Sears MR, Rea HH, de Boer B, et al. Accuracy of certification of deaths due to asthma: a national study. *Am J Epidemiol.* 1986;124:1004–1011.
26. Barger LW, Vollmer WM, Felt RW, et al. Further investigations into the recent increase in asthma death rates: a review of 41 asthma deaths in Oregon in 1982. *Ann Allergy.* 1988;60:30–39.
27. Robinson WS. Ecological correlations and the behavior of individuals. *Am Sociological Rev.* 1950;15:351–357.
28. Kerner JF, Andrews H, Zauber Z, Struening E. Geographically-based cancer control: methods for targeting and evaluating interventions on defined populations. *J Clin Epidemiol.* 1988;41:543–553.

Diversity of Tick Species Biting Humans in an Emerging Area for Lyme Disease

Robert P. Smith, Jr, MD, MPH, Eleanor H. Lacombe, MT(ASCP), Peter W. Rand, MD, and Richard Dearborn

ABSTRACT

Background. Although most tick bites in humans in areas of the northeastern United States in which Lyme disease is highly endemic are due to *Ixodes dammini*, no study documents the frequency of *I. dammini* bites in low-prevalence or emerging areas for Lyme disease. Data on the proportion of tick bites in humans that are due to *I. dammini* in a region may have implications for public health policy and clinical management.

Methods. A statewide survey of the tick species that parasitized humans in Maine was conducted during 1989 and 1990. Tick submissions from throughout the state were elicited through media announcements. All ticks that had been removed from humans were identified, and data were collected that included bite seasonality and geography and demographics of tick bite victims.

Results. Of 709 ticks submitted, only 17% were *I. dammini. Ixodes cookei*, a vector for Powassan encephalitis, accounted for 34% of bites, and *Dermacentor variabilis* accounted for 45%. Other tick species were occasionally implicated.

Conclusions. The likelihood that a tick bite was due to *I. dammini* was lower in Maine than in areas in the northeastern United States in which Lyme disease is highly endemic. Other tick vectors, associated with diseases other than Lyme disease, were more frequently implicated. Regional tick bite surveys may prove useful in assessing the risk of Lyme disease following a tick bite. (*Am J Public Health.* 1992;82:66–69)

Introduction

In areas in which Lyme disease is highly endemic, the northern deer tick, *Ixodes dammini*, accounts for up to three fourths of tick bites in humans.[1,2] As many as 30% to 70% of nonlarval deer ticks in these areas may be infected with *Borrelia burgdorferi*.[3–6] Because unrecognized Lyme disease can have severe consequences, many physicians in high-prevalence areas routinely treat patients with a recent history of a tick bite with oral antibiotics in an attempt to prevent the development of infection. One double-blind placebo-controlled study of this practice failed to demonstrate a clear benefit with treatment, even though all study participants were bitten by *I. dammini* and tick infection rates were 30%.[7] Because of public fear of Lyme disease and a lack of regional data on the likelihood of tick bites by *I. dammini*, physicians in regions with any Lyme disease risk may consider antibiotic treatment of all tick bites. But in areas recently colonized by *I. dammini*, or at the edge of its range, the risk of Lyme disease following tick bites may be negligible because of limited distribution and a low population density of *I. dammini*.

From 1985, when the deer tick was first identified in Maine,[8] through 1989 there have been only 14 reported cases of Lyme disease acquired in Maine that meet the Centers for Disease Control's case definition. Most cases have been reported from two counties in southwest coastal Maine. Although deer ticks have been found in many coastal areas of the state, only these two counties were considered endemic areas for Lyme disease in 1989 by the Centers for Disease Control. This study characterizes tick species that parasitize humans in Maine, an emerging area of Lyme disease with a low prevalence of infection.

Materials and Methods

As part of a multifaceted study of *I. dammini* distribution in Maine, the Research Laboratory at the Maine Medical Center and the Maine Department of Conservation Insect and Disease Laboratory collaborated on a study of human parasitism by ticks during 1989 and 1990. Requests for ticks were publicized in brochures, public health notices, and the media. Individuals bitten by ticks in Maine were encouraged to save the ticks in alcohol and submit them for identification to either laboratory. The bite victim's name, age, and address; the data and locale of the bite; and any associated symptoms were requested. All ticks received were examined under a dissecting microscope and identified according to standard keys and taxonomic references.[9–13] All ticks submitted to state agencies were identified by the state entomologist; specimens received at the Maine Medical Center Research Laboratory were identified by research personnel with training and experience in tick identification (E.H.L., R.P.S.). Unusual specimens were submit-

Robert P. Smith, Jr, is with the Department of Medicine and Eleanor H. Lacombe and Peter W. Rand are with the Research Department, all at Maine Medical Center, Portland, Me. Richard Dearborn is with the Maine Forest Service, Augusta, Me.

Requests for reprints should be sent to Robert P. Smith, Jr, MD, Maine Medical Center, Research Department, 22 Bramhall Street, Portland, ME 04102.

This paper was submitted to the journal July 9, 1990, and accepted with revisions May 7, 1991.

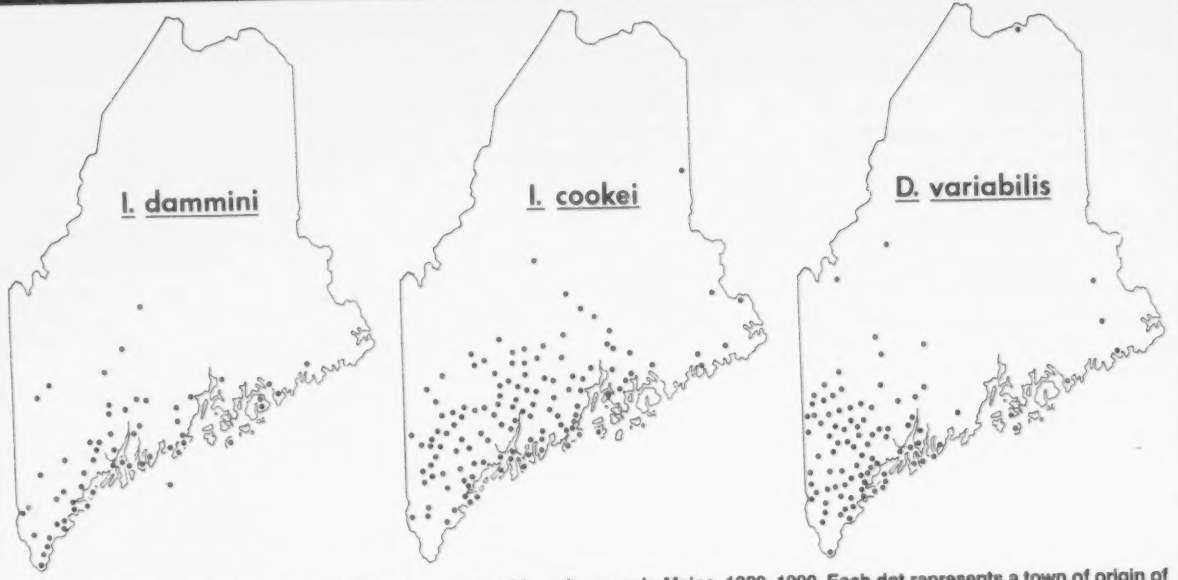


FIGURE 1—Distribution of common tick species removed from humans in Maine, 1989–1990. Each dot represents a town of origin of identified ticks.

ted to Dr J. E. Keirans, Smithsonian Institution, for confirmation of species identification.

Results

A total of 709 ticks that had been removed from 646 humans in Maine were submitted for identification. Ninety percent of all *Ixodes* ticks had been attached at time of removal. Although submitted ticks were received from residents of all of Maine's counties, few specimens were sent from northern areas (Figure 1). In the two known endemic counties, only 13% of ticks removed from humans were *I. dammini.* Of all *Ixodes* ticks submitted from these counties, 43.7% were *I. dammini*, compared with 26.3% in the remaining counties.

Tick species and stages are listed in the Table. Distribution of biting ticks of common species is shown in Figure 1. Seasonal distribution of tick bites by species is illustrated in Figure 2. The ages of tick bite victims ranged from 1 month to 89 years. Fifty percent of the *Ixodes cookei* ticks were removed from children under 16 years old, as opposed to 23% of *I. dammini* (odds ratio, 3.38 [95% confidence interval, CI, 2.05–5.57], $P < .001$) and 28% of *Dermacentor variabilis* (odds ratio, 2.52 [95% CI, 1.74–3.65], $P < .001$).

Ticks Removed from Humans in Maine, 1989–1990

	Hosts (n)	Ticks (n)	Larvae	Nymphs	Adults
I. dammini	120	123	2	19	102
I. cookei	217	230	24	179	27
I. muris	8	8	—	1	7
I. marxi	19	20	—	15	5
D. variabilis	276	320	—	2	318
D. albipictus	6	8	5	3	—
Total	646	709	31	219	459

Discussion

Tick surveys of deer and small rodents in Maine document the presence of *I. dammini* in widely dispersed coastal sites.[14,15] In contrast to reports from areas in which Lyme disease is highly endemic, however, this study demonstrates that *I. dammini* accounts for only a minority of tick bites in humans, even in the two Maine counties recognized as endemic in 1989. Sampling bias due to a low level of vigilance for small ticks could underestimate the abundance of *I. dammini* in an emerging area for Lyme disease, and such a bias is suggested by a low ratio of nymphal to adult *I. dammini* in this study.[1,2] However, the most common *Ixodes* tick species in our survey, *I. cookei*, is similar in size to *I. dammini*, and numerous specimens of the nymphal and larval stages of this tick were submitted.

Given the small proportion of tick bites due to *I. dammini*, the potential for exposure to Lyme disease from a tick bite in Maine is very low. As disease transmission requires attachment for 24 hours or more, many bites by disease-carrying ticks will not result in infection if the tick is removed.[16,17] One study in an area in which Lyme disease is highly endemic did not show a clear benefit from antibiotic treatment of deer tick bites, with only one of 30 untreated bites resulting in disease.[7] A second epidemiologic study of Lyme disease in outdoor workers, however, suggested a protective effect from incidental antibiotic use during summer months.[18] Our study demonstrates striking regional differences in the likelihood that a tick bite is due to a deer tick and supports the need for caution in applying results from studies in highly endemic areas to regions of lower prevalence of Lyme disease.

In contrast to previous surveys in the northeastern United States, this study documents frequent human parasitism by *I. cookei*, the "woodchuck tick." This small tick feeds on medium-sized mammals and cannot be distinguished from *I. dammini* without the aid of a dissecting microscope.[19–21] It has been reported as a

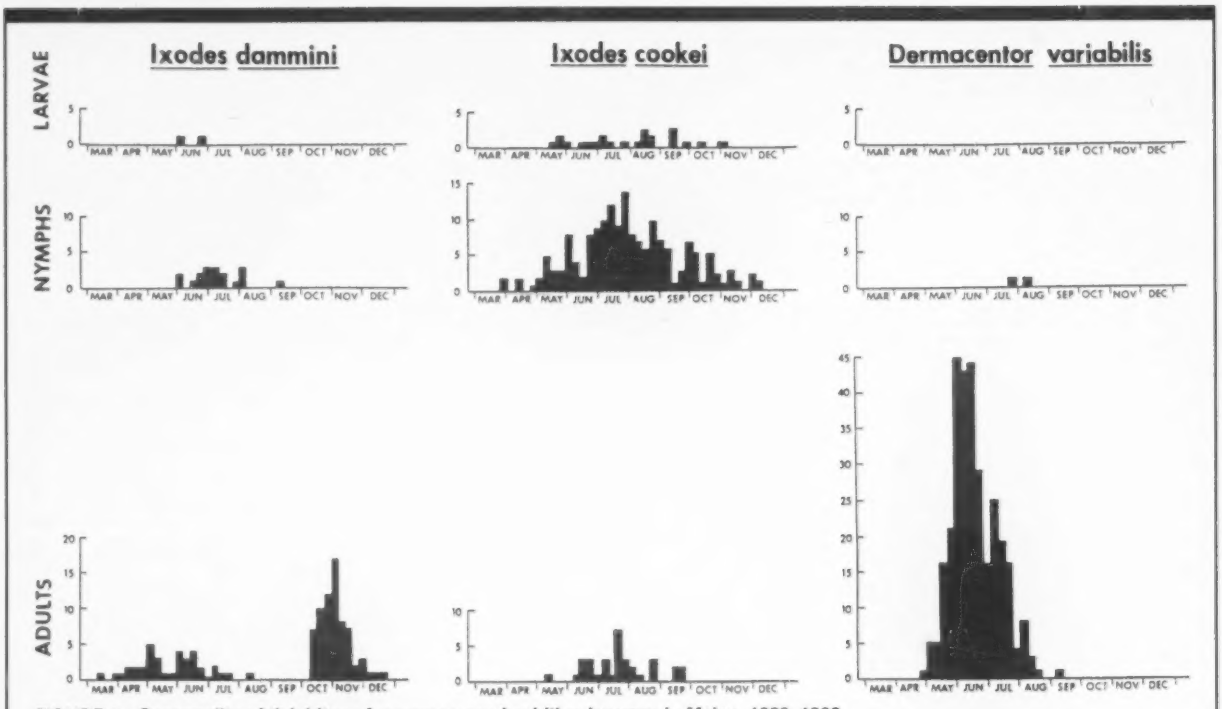


FIGURE 2—Seasonality of tick bites of common species biting humans in Maine, 1989–1990.

common cause of tick bites in Ontario, but has rarely been found to parasitize humans in other studies.[1,2,22] Although *B. burgdorferi* infection of *I. cookei* has been documented in a few tick specimens,[23] the role of this tick as a possible vector of Lyme disease has not been well studied. Unlike *I. dammini*, in its immature stages *I. cookei* does not parasitize mice.[20] *I. cookei* is a vector of Powassan encephalitis, a Group B arbovirus disease reported from upstate New York and Canada,[24,25] but not from Maine. Powassan encephalitis, although rarely diagnosed, can cause meningoencephalitis and radiculoneuritis with a spectrum of symptoms that could be confused with neurological complications of Lyme disease.[25–27]

Other small *Ixodes* species (*I. muris, I. marxi*) occasionally parasitized humans in Maine. *I. marxi* is another known vector of Powassan encephalitis.[21] The vector competence of both of these species for *B. burgdorferi* is unknown. *I. muris* feeds primarily on small rodents and could theoretically be an occasional vector of Lyme disease to humans.[28] Another mouse-feeding tick, *I. angustus*, has been cited as a vector of Lyme disease in one case report from the northwestern United States.[29] Despite its abundance on mice and red-backed voles in Maine, no examples of human parasitism by this tick were documented. *I. angustus* appears to be an unimportant vector of Lyme disease to humans in the northeastern United States.

D. variabilis, which accounted for 45% of bites, is widely distributed in southwestern Maine.[30] *D. variabilis* ticks may occasionally be infected with *B. burgdorferi*, but laboratory studies indicate that they are an ineffective vector of the disease.[31] *D. albipictus*, which is host specific for deer and moose, rarely bites humans but was in fact detached from one deer processor. It is not known to be a vector of disease to man.

Marked differences in the seasonal activity of tick species are illustrated by this study. In Maine, as in other areas of the northeastern United States, humans were parasitized by *I. dammini* nymphs in early summer and by adults in late fall.[32] The presence of biting *I. dammini* adults in late fall is of possible epidemiologic importance and may account for a bimodal peak of Lyme disease incidence reported in two previous studies.[5,33] Also of interest was the sporadic presence of biting adult deer ticks throughout the spring and summer months. These adult deer ticks, when engorged, may overlap in size with unengorged adult *D. variabilis*, for which they were sometimes mistaken by physicians. *I. cookei* infestation occurred throughout the summer, with no relationship between tick stages and month of bite. Ticks removed from humans in August and September, when active *I. dammini* are predominantly larvae, were almost exclusively *I. cookei*.

To the extent that they demonstrate changes in the abundance of *I. dammini* in emerging areas over time, surveys of human parasitism by ticks provide essential information for the development of regional public health policy regarding Lyme disease. In addition, tick identification services may highlight unexpected vectors, such as *I. cookei*, that may be the cause of other infectious diseases that could be mistaken for Lyme disease. However, because the diversity and abundance of tick species may vary markedly over small areas, depending on the presence or absence of appropriate host animals, habitat type, and other factors, the risk of exposure to Lyme disease from tick bites may vary substantially at different sites in the same region.[36]

This study underscores the need for caution in the identification of tick species. Although the majority of "small" ticks biting people in some areas of the northeastern United States may be *I. dammini*, this is clearly not the case in Maine. The identification of different tick species in the *Ixodes* genus depends on the recognition of subtle morphological differences and

requires training and experience. A lack of government and academic support for arthropod biosystematics has contributed to the paucity of entomologists or others with these skills.[37] Effective regional tick surveys, therefore, may require cooperative arrangements between state health personnel and entomologists at colleges and universities, or others with training in tick identification. □

Acknowledgments

Funding was provided by the Davis Family Foundation, the Simmons Foundation, and L. L. Bean, Inc.

The authors gratefully acknowledge Drs J. E. Keirans and Richard Robbins at the Smithsonian Institution's Research Division for their assistance with tick identification. We also acknowledge the statistical review of our data by Dr W. Douglas Thompson, University of Southern Maine.

References

1. Falco RC, Fish D. Ticks parasitizing humans in a Lyme disease endemic area of southern New York State. *Am J Epidemiol.* 1988;128:1146–1152.
2. Magnarelli LA, Anderson JF. Infected ticks feeding on persons in areas endemic for Lyme disease and Rocky Mountain spotted fever. *J Infect Dis.* 1988;160:729–730.
3. Magnarelli LA, Anderson JF, Apperson CS, et al. Spirochetes in ticks and antibodies to *Borrelia burgdorferi* in white-tailed deer from Connecticut, New York State, and North Carolina. *J Wildlife Dis.* 1986;22:178–188.
4. Anderson JF, Magnarelli LA, Burgdorfer W, Barbour AG. Spirochetes in *Ixodes dammini* and mammals from Connecticut. *Am J Trop Med Hyg.* 1983;32:818–824.
5. Lastavica CC, Wilson MM, Berardi VP, Spielman A, Deblinger RD. Rapid emergence of a focal epidemic of Lyme disease in coastal Massachusetts. *New Engl J Med.* 1989;320:133–137.
6. Anderson JF, Duray PH, Magnarelli LA. *Borrelia burgdorferi* and *Ixodes dammini* prevalent in the greater Philadelphia area. *J Infect Dis.* 1990;161:811–812.
7. Costello CA, Steere AC, Pinkerton RE, et al. A prospective study of tick bites in an endemic area for Lyme disease. *J Infec Dis.* 1990;159:136–138.
8. Anderson JF, Magnarelli LA. *Ixodes dammini* and *Borrelia burgdorferi* in northern New England and upstate New York. *J Parasitol.* 1987;73:419–421.
9. Cooley RA, Kohls GM. The genus *Ixodes* in North America. *Natl Inst Health Bull.* 1945;184:1–246.
10. Keirans JE, Clifford CM. The genus *Ixodes* in the U.S.: a scanning electron microscope study and key to adults. *J Med Entomol.* 1978;2(suppl):1–149.
11. Spielman A, Clifford CM, Piesman J, Corwin MD. Human babesiosis on Nantucket Island, U.S.A.: Description of a vector, *Ixodes* (Ixodes) dammini n. sp. (Acari: Ixodidae). *J Med Entomol.* 1979;15:218–234.
12. Clifford CM, Anastos G, Elbl A. The larval *Ixodes* ticks of the eastern United States (Acarina-Ixodidae). *Entomol Soc Am.* 1961;2:213–237.
13. Keirans JE, Litwak TR. Pictorial key to the adults of hard ticks, family Ixodidae (Ixodida: Ixodoidea), east of the Mississippi. *J Med Entomol.* 1989;26:435–448.
14. Smith RP, Rand PW, Lacombe EH. Potential for Lyme disease in Maine—deer survey for *Ixodes dammini*, the tick vector. *Am J Public Health.* 1990;80:333–335.
15. Ginsberg HS, Ewing C. Deer ticks, *Ixodes dammini* (Acari: Ixodidae) and Lyme disease spirochetes, *Borrelia burgdorferi*, in Maine. *J Med Entomol.* 1988;25:303–304.
16. Piesman J, Mather TN, Sinsky RJ, et al. Duration of tick attachment and *Borrelia burgdorferi* transmission. *J Clin Microbiol.* 1987;25:557–558.
17. Gerber MA, Shapiro ED, Persing DH, et al. Risk of developing Lyme disease after a deer tick bite. Fourth International Conference on Lyme Borreliosis, Stockholm, Sweden. Abstract.
18. Schwartz BS, Goldstein MD. Lyme disease in outdoor workers: risk factors, preventive measures, and tick removal methods. *Am J Epidem.* 1990;131:877–885.
19. Farkas MJ, Surgeoner G. Incidence of *Ixodes cookei* (Acari: Ixodidae) on groundhogs, *Marmota monax*, southwestern Ontario. *Proc Entomol Soc Ontario.* 1990; 121:105–110.
20. Ko RC. Biology of *Ixodes cookei* Packard (Ixodidae) of groundhogs (*Marmota monax* Erxleben). *Can J Zool.* 1972;50: 432–436.
21. Main AJ, Corey AB, Downs WG. Powassan virus in *Ixodes cookei* and Mustelidae in New England. *J Wildlife Dis.* 1979;15:585–591.
22. Scholten TH. Human tick infestations in Ontario: findings at the Toronto Public Health Laboratory, 1967–1977. *Can J Public Health.* 1977;68:494–496.
23. Hall JE, Anrine JW, Gais RD, et al. Parasitization of humans in West Virginia by *Ixodes cookei* (Acari: Ixodidae), a potential vector of Lyme borreliosis. *J Med Entomol.* 1991;28:186–189.
24. Artsob H, Spence L, Surgeoner G, et al. Isolation of *Francisella tularensis* and Powassan virus from ticks (Acari: Ixodidae) in Ontario, Canada. *J Med Entomol.* 1984;21:165–168.
25. Smith R, Woodall JP, Whitney E, et al. Powassan virus infection. A report of three human cases of encephalitis. *Am J Dis Child.* 1974;127:691–693.
26. Embil JA, Camfield P, Artsob H, et al. Powassan virus encephalitis resembling herpes simplex encephalitis. *Arch Intern Med.* 1983;143:341–344.
27. Conway H, Rossier E, Spencer L, et al. Powassan virus encephalitis with shoulder girdle involvement. *Can Dis Weekly Rep.* 1976;2:85–87.
28. Smith CN. Biology of *Ixodes muris* Bishopp and South (Ixodidae). *Ann Entomol Soc Am.* 1944;37:221.
29. Damrow T, Bowen GS, Lakat MF, et al. Is *Ixodes* (Ixodiopsis) *angustus* a vector of Lyme disease in Washington state? *West J Med.* 1989;150:580–582.
30. State of Maine. *Twenty-eighth Biennial Report of the Forest Commissioner, A. D. Nutting, 1951–1952.*
31. Piesman J, Sinsky RJ. Ability of *Ixodes scapularis, Dermacentor variabilis*, and *Amblyomona americanum* (Acari: Ixodidae) to acquire, maintain, and transmit Lyme disease spirochetes (*Borrelia burgdorferi*). *J Med Entomol.* 1988;25:336–339.
32. Wilson MC, Spielman A. Seasonal activity of immature *Ixodes dammini* (Acari: Ixodidae). *J Med Entomol.* 1988;22:408–414.
33. Schulze TL, Bower GS, Lakat MF, et al. The role of adult *Ixodes dammini* (Acari: Ixodidae) in the transmission of Lyme disease in New Jersey, USA. *J Med Entomol.* 1985;22:88–93.
34. Spielman A, Wilson ML, Levine JF, et al. Ecology of *Ixodes dammini*-borne human babesiosis and Lyme disease. *Ann Rev Entomol.* 1985;30:439–460.
35. Ginsberg H, Ewing CP. Habitat distribution of *Ixodes dammini* (Acari: Ixodidae) and Lyme disease spirochetes on Fire Island, New York. *J Med Entomol.* 1989;26:183–189.
36. Falco RC, Fish D. Potential for exposure to tick bites in recreational parks in a Lyme disease endemic area. *Am J Public Health.* 1989;79:12–15.
37. Oliver JH. Crisis in biosystematics of arthropods. *Science.* 1988;240:967.

Prevalence of Human T-Lymphotropic Virus in Civilian Applicants for the United States Armed Forces

Chester R. Roberts, PhD, Donald R. Fipps, MS, John F. Brundage, MD, Stephen E. Wright, MS, Mary Goldenbaum, MS, Steven S. Alexander, PhD, and Donald S. Burke, MD

ABSTRACT

Background. The wide range in human T-lymphotropic virus (HTLV) seroprevalences reported worldwide has made estimates of seroprevalence difficult in unique populations. In this study, the seropositivity of young adult civilian applicants for the US Armed Forces was determined.

Methods. Serum samples from nine geographic regions were screened by an Enzyme-Linked Immunosorbent Assay (ELISA), and repeatedly reactive samples were further tested by Western blot and radioimmunoprecipitation. Specimens were scored as positive when antibody to *gag* (p24) and *env* (gp46 or gp68) were detected.

Results. Of the 43 750 samples analyzed, 18 were positive for HTLV antibodies. Ten (55%) were from males and eight (45%) were from females. Nine (90%) of the males and seven (87.5%) of the females were Black. Twelve of the positive samples (66.6%) were from the New York City region, which represented only 18.8% of the sample population.

Conclusions. The overall HTLV seroprevalence of civilian applicants for the US Armed forces was 0.41 per 1000. This was higher than the seroprevalence reported for volunteer blood donors. (*Am J Public Health.* 1992;82:70–73)

Introduction

The human T-lymphotropic viruses type I (HTLV-I) and type II (HTLV-II) are retroviruses that are being detected in a wide variety of populations. It is likely that the virus is transmitted in several ways, including through contaminated blood products, exposure to contaminated devices used for illicit intravenous drug use, colostrum ingestion, and sexual contact. Two well-characterized, albeit rare, clinical conditions that result from HTLV-I infection are adult T-cell leukemia and tropical spastic paraparesis. Acute development of these conditions has not been documented; rather, clinically overt signs of disease appear to develop many years after infection.[1–3] Although seroconversion related to blood transfusion is well documented, there have been no reports of clinical manifestations or development of ATL directly associated with exposure to infected blood or blood products.[4] HTLV-II has not been firmly established with any clinical syndromes.

The prevalence of HTLV antibody has been studied in many different areas of the world. The southern islands of Japan appear to have some of the highest prevalence rates. Surveys have reported the following high prevalences among adults in these areas of Japan: Nagasaki, 37%; Kagoshima, 24%; and Okinawa, between 10% and 51%.[5–7]

Hawaiian-Japanese emigrants from high-prevalence areas have a reported prevalence rate of 20%.[8] In the Caribbean, the rate in Martinique blood donors was reported as 4.0%.[9] Studies of intravenous drug abusers who shared needles showed prevalences that varied from 9.0% in New York[7] to 64.6% in Rome.[10] Prevalence studies in presumed low-risk populations in the United States have shown a wide range of HTLV seropositivity: for example, seroprevalence rates were 2.1% in Georgia[7,11]; 0.5% in Florida[11]; 1% to 1.4% in Burlington, Vt[12,13]; 0.5% in Washington, DC[7]; 4.2% in Houston, Tex; and 3.8% in Birmingham, Ala, in blood donors over 30 years old.[12] An American Red Cross study of 39 898 blood donors from eight geographical areas in the United States showed a prevalence rate of 0.025%.[14,15] All of these blood donors were asymptomatic.

The wide range of HTLV seropositivity reported in these studies made it important to investigate the prevalence of detectable antibody to HTLV in civilian applicants for the Armed Services of the United States of America. These data would be critical to enable health care planners and providers to assess the impact of HTLV seroprevalence on health care operations.

Chester R. Roberts, Donald R. Fipps, and Donald S. Burke are with the Division of Retrovirology and John F. Brundage and Mary Goldenbaum are with the Division of Preventive Medicine, all at the Walter Reed Army Institute of Research, Washington, DC. Stephen E. Wright is with the Operations, Plans, and Training Directorate, Headquarters, US Military Entrance Processing Command, Chicago, Ill. Steven S. Alexander is with Biotech Research Laboratories, Rockville, Md.

Requests for reprints should be sent to Chester R. Roberts, PhD, Division of Retrovirology, Walter Reed Army Institute of Research, 13 Taft Court, Suite 201, Rockville, MD 20850.

This paper was submitted to the journal November 1, 1990, and accepted with revisions June 10, 1991.

Locations from which Sera were Collected for HTLV-I/II Antibody Seroprevalence Tests

Geographic Area	Number Sera	Number HTLV-Positive
New York, NY	8239	12
Baltimore, Md	7925	3
Cincinnati, Ohio	3249	0
New Orleans, La	5602	0
San Juan, Puerto Rico	2694	1
Fargo, ND	1800	0
Honolulu, Hawaii	1379	0
Miami, Fla	4196	1
Oakland, Calif	8666	1
Total	43,750	18

Methods

Population

Specimens for this study were a collected subset of samples of serum screened for human immunodeficiency virus (HIV). The specimens were collected at nine selected geographic regions from civilian applicants for the Armed Services of the United States between March 1988 and August 1989 (see Table). For this survey, specific identifier information (i.e., name and Social Security number) were unlinked from the specimen in order to ensure confidentiality.

Since HTLV-I/II and HIV-1 have similar modes of transmission, it was assumed that HTLV-I/II prevalence would mirror HIV-1 prevalence.[16] Therefore, the New York City, Baltimore, and Oakland regions were selected because of the high prevalence of HIV-1 in these geographic areas. Cincinnati, Ohio, was studied as an intermediate control with a 0.44 per thousand HIV-1 prevalence rate. Fargo, ND, was used as a negative control, because during the past 2 years, not a single Western-blot–positive HIV-1-infected individual has been processed through that region. The New Orleans region was selected because of reports that intravenous drug abusers in that community have a higher prevalence of HTLV-II and are not coinfected with HIV-1.[17] Honolulu was selected to determine whether applicants have increased prevalence rates because of Japanese influence. Miami and San Juan, Puerto Rico, were selected to determine whether applicants being processed from the Caribbean basin have an increased prevalence rate.

Serology

The screening test used to detect antibody to HTLV was a qualitative enzyme-linked immunosorbent assay (ELISA), obtained from the DuPont Company, Wilmington, Del. All repeatedly reactive samples were tested using the Western blot method (Biotech Research Laboratories, Inc, Rockville, Md). A blot clear of any bands, or containing bands seen where viral proteins do not migrate (nonspecific banding), was considered nonreactive. Blots were considered reactive when a gp46 (*env*) band and a p24 (*gag*) band were seen. The presence of viral bands in any combination other than the aforementioned pattern was considered nondiagnostic. All repeatedly reactive specimens were also subjected to a radioimmunoprecipitation assay (RIPA) performed by Biotech Research Laboratories, Rockville, Md. The HTLV antigen used in the RIPA was prepared from the MT-2 cell line. The presence of a band at 68 000 Da was considered reactive for antibody specific to the *env* protein of HTLV-1. In addition, the specimens were subjected to testing for the presence of antibodies to the recombinant HTLV-1 envelope protein known as 21E.[18] This included testing by an ELISA obtained from Cambridge Bioscience Corporation, Worcester, Mass, and by a Western blot performed by Biotech Research Laboratories, Rockville, Md.

Specimens were scored as positive when antibody to *gag* (p24) and *env* (gp46 or gp68) could be demonstrated by Western blot alone or in combination with RIPA. Due to the cross-reactivity of anti-HTLV-II antibodies in all of the above serological methods, positive specimens may be due to either HTLV-I or HTLV-II.

Results

In the overall study population (see Appendix), 82.6% (36 138) were male and 17.4% (7612) were female. Approximately 53% of the study group were less than 21 years old. The majority of the applicants (53.9%) were White, followed by Black (26.7%) and Hispanic (12.1%). Several applicants (7.3%) selected "other" as their ethnicity. Although the study applicants were processed through the nine geographic regions sampled, they represented some 40 different states and US territories.

A total of 43 750 specimens were screened during the study period, and 285 were found to be repeatedly reactive for a total ELISA repeat reactive rate of 6.51 per thousand. Prevalences of repeat reactivity varied among study sites, from 0.18 (Honolulu, Hawaii) to 1.51 (New York City, NY) per thousand. Of the 285 repeat reactives, 119 (41.8%) were nonreactive, 158 (55.4%) were nondiagnostic, and 8 (2.8%) were reactive by Western blot. All these specimens were also tested by RIPA. The RIPA was able to demonstrate *env* protein (gp68) on all eight of the reactive blot samples, 10 of the nondiagnostic samples that also had a p24, and none of the nonreactive blot samples. Therefore, using the criteria discussed above, 18 specimens were confirmed as positive. The overall prevalence for detectable antibody to HTLV-I/II in this sample of civilian applicants for the US Armed Forces was thus determined to be 0.41 per thousand. Of the positive applicants, 10 (55%) were male and 8 (44.5%) were female. Nine (90%) of the males and 7 (87.5%) of the females were Black. The overall prevalence among Blacks was 1.37 per 1000. One of the Black females was also HIV infected. The youngest male was 18 and the oldest was 43, with the average age being 28.0 and the median age being 22. The youngest female was 20 and the oldest 47, with the average age being 28.5 and the median age being 24. Twelve of the positives (66.6%) were from the New York City region, which represented only 18.8% of the sample population. Since 895 000 individuals volunteered for military duty during the study period, if we assume that this study is representative of those volunteering for military duty it can be estimated that over 260 HTLV-I/II infected individuals entered the military service between March 1988 and August 1989.

Discussion

In a recent report on 39 898 American Red Cross volunteer blood donors, Fang et al. reported an initial reactive rate to HTLV-I of 0.44% and a repeat reactive rate of 0.17%.[14] Our repeat reactive rate of 0.65% among civilian applicants for the military was 3.8 times higher than that among volunteer blood donors. This may have been due to the different screening ELISAs employed. However, the prevalence of confirmed positives (0.41 per thousand) among civilian applicants was 1.6 times higher than that among volunteer blood donors. This finding starts to quantify the view that the prevalence of

HTLV-I/II in blood donors significantly underestimates that in the general population.

Determining HTLV-I/II antibody status will help define the epidemiology of the disease. More importantly, it may prevent HTLV-positive individuals from passing the virus to sexual partners and to newborn infants through infected mother's breast milk, and will exclude infected blood and blood products from blood bank inventories as well. For example, in Japan, the seroconversion rates to HTLV-I due to blood transfusions were reduced from 53.6% to 0.9% by removing seropositive blood products from the inventory.[19] Although detecting blood infected with HTLV-I or any retrovirus and removing it from the inventory is for the protection of the intended recipient, the notification, counseling, and access to health care benefits of the infected individual are equally important. When the demographic data of the 18 positive samples are compared with those of the overall study population, some surprising contrasts are apparent (see Appendix). Although only 17.4% of the study population were female, almost half (44.5%) of the HTLV-positive individuals were female. The prevalence among females was almost four times greater than that among males. Black applicants represented 26.7% of the study population but 88.8% of the HTLV-positive samples. The prevalence among Blacks was more than 30 times greater than that among whites. These relationships are based on unstable estimates and are confounded by geographic region. Still, they may provide some insight regarding the current distribution of risk among young adults in the United States.

For a number of reasons, the population in this study cannot be considered an unbiased sample of the general US population. The study subjects were not geographically representative, either by region or area (urban vs rural), of the US population. The general population in the United States is not 82.6% male, or 26% Black, nor is 53% of the population less than 21 years old. Several studies have shown that the prevalence of HTLV-I increases with age of population.[6,11,20–22] As the majority of the study population was under 21 years old, the true prevalence of infection in general is likely to exceed that reported here. In addition, this sample excludes individuals who have been convicted of a felony and those who are overt users of illicit drugs. This sample also probably underrepresents those who know or suspect that they are infected with HIV-1 (e.g., those identified as HIV-1-positive as blood donors through alternative test sites). For these reasons, the prevalences of antibodies found in this survey almost certainly underestimate the actual prevalences of HTLV-I/II infection in the general adult population from corresponding regions of the United States. However, from these data, one can estimate that the total HTLV-I/II-positive population currently entering the Armed Forces would be approximately 230 per year, providing recruiting numbers remain constant at approximately 564 000 per year. For health care planners and providers, HTLV infection would have a minimal impact on health care resources in comparison with other diseases of military importance. □

References

1. Shimoyama M. Adult T-cell leukemia/lymphoma and its clinical subtypes from viewpoints of viral etiology. *Curr Top Microbiol Immunol*. 1985;115:113–126.
2. Morishima Y, Ohya KI, Ueda R, Fukuda T. Detection of adult T-cell leukemia virus (ATLV) bearing lymphocytes in concentrated red blood cells derived from ATL associated antibody (ATLA-Ab) positive donors. *Vox Sang*. 1986;50:212–215.
3. Takatuki K, Yamaguchi K, Kawano F, Nishimura H, Seiki M, Yoshida M. Clinical aspects of adult T-cell leukemia/lymphoma. *Curr Top Microbiol Immunol*. 1985;115:89–98.
4. Okochi K, Sato H. Transmission of adult T-cell leukemia virus (HTLV-I) through blood transfusion and its prevention. *AIDS Research*. 1986;2(2 Pt 1):S157–S161.
5. Hinuma Y, Komoda H, Kondo T, et al. Antibodies to adult T-cell leukemia-virus-associated antigen (ATLA) in sera from patients with ATL and controls in Japan: a nationwide seroepidemiological study. *Int J Cancer*. 1982;29:631–635.
6. Kajiyama W, Kashiwagi S, Ikematsu H, Hayashi J, Nomura H, Okochi K. Intrafamilial transmission of adult T-cell leukemia virus. *J Infect Dis*. 1986;154:851–857.
7. Blattner WA, Blayney DW, Robert-Guroff M, et al. Epidemiology of human T-cell leukemia/lymphoma virus. *J Infect Dis*. 1983;147:406–416.
8. Blattner WA, Nomura A, Clark JW, Ho GYF, Gallo R, Robert-Guroff M. Modes of transmission and evidence for viral latency from studies of human T-cell lymphotropic virus type I in Japanese migrant populations in Hawaii. *Proc Natl Acad Sci USA*. 1986;83:4895–4898.
9. Gessain A, Barin F, Vernant JC, et al. Antibodies to human T-lymphotropic virus type I in patients with tropical spastic paraparesis. *Lancet*. 1985;ii:407–410.
10. Rezza G, Titti F, Rossi GB, Verani P, Menniti-Ippolito F, Oliva C. Sex as risk factor in HTLV-I spread among intravenous drug abusers. *Lancet*. 1988;i:713.
11. Blayney DW, Blattner WA, Robert-Guroff M, et al. The human T-cell leukemia-lymphoma virus in southeastern United States. *JAMA*. 1983;250:1048–1053.
12. Blattner WA, Gallo RC. Human T-cell leukemia/lymphoma viruses: clinical and epidemiological features. *Curr Top Microbiol Immunol*. 1985;115:68–88.
13. Minamoto GY, Gold JWM, Scheinberg DA, et al. Infection with human T-cell leukemia virus type I in patients with leukemia. *N Engl J Med*. 1988;318:219–222.
14. Fang CT, Williams AE, Sandler SG, Slamon DJ, Poiesz BJ. Detection of antibodies to human T-lymphotropic virus type I (HTLV-I). *Transfusion*. 1988;28:179–183.
15. Williams AE, Fang CT, Salmon DJ, et al. Seroprevalence and epidemiological correlates of HTLV-I infection in U.S. blood donors. *Science*. 1988;240:643–646.
16. Burke DS, Brundage JF, Herbold JR, et al. Human immunodeficiency virus infections among civilian applicants for United States military service, October 1985 to March 1986. *N Engl J Med*. 1987;317:132–136.
17. Wiktor S, Alexander S, Shaw G, et al. Distinguishing between HTLV-I and HTLV-II by Western blot. *Lancet*. 1989;335:1533.
18. Samuel K, Lautenberger J, Jorcyk C, Josephs S, Wong-Staal F, Papas T. Diagnostic potential for human malignancies of bacterially produced HTLV-I envelope protein. *Science*. 1984;226:1094–1097.
19. Kamihira S, Nakasima S, Oyakawa Y, et al. Transmission of human T-cell lymphotropic type I by blood transfusion before and after mass screening of sera from seropositive donors. *Vox Sang*. 1987;52:43–44.
20. Clark J, Saxinger C, Gibbs WN, et al. Seroepidemiologic studies of human T-cell leukemia/lymphoma virus type I in Jamaica. *Int J Cancer*. 1985:36:37–41.
21. Botha DW, Jones M, Klerk WA, Yamamoto N. Distribution and possible spread of human T-cell leukemia virus type I in human communities in northern and eastern Transvall. *S Afr Med J*. 1985;67:668–671.
22. Moss A. HTLV-I prevalence in IV drug users. Presented at the first Abbott HTLV-I Technical Conference; May 3, 1988; Abbott Park, Ill.

APPENDIX--Demographics of HTLV-I/II Antibody Positive Applicants

A. Geographic Area	Gender		Race	Age	
	Male	Female		Male	Female
New York, NY	7	5	Black	18, 20, 22, 30, 35, 35, 43	20, 20, 21, 24, 33
Baltimore, Md	1	1	Black	20	47
		1	White		36
San Juan, Puerto Rico	0	1	Black	—	27
Miami, Fla	1	0	Black	18	—
Oakland, Calif	1	0	White	39	—

B. Gender	No. Positive/Total Tested	Prevalence (95% CI, Poisson)
Male	10/36,138	0.28/1000 (0.13, 0.51)
Female	8/7612	1.05/1000 (0.45, 2.07)

C. Race/Ethnicity	No. Positive/Total Tested	Prevalence (95% CI, Poisson)
White	2/23,581	0.08/1000 (0.001, 0.24)
Black	16/11,681	1.37/1000 (0.85, 2.33)
Hispanic	0/5294	0.0/1000 (0.0, 0.70)
Other	0/3194	0.0/1000 (0.0, 1.15)

Correlates of Using Mental Health Services: Implications of Using Alternative Definitions

David Mechanic, PhD, Ronald J. Angel, PhD, and Lorraine Davies, MA

ABSTRACT

Background. Studies using varying definitions of mental health visits result in widely divergent estimates. This analysis examines the stability of a predictive model using three varying definitions based on provider type, diagnosis, psychotherapy, and psychotropic medication use.

Methods. Interview and questionnaire data and claims records from the RAND Health Insurance Experiment are used to test these models among 3138 enrollees.

Results. Estimates of visits, and factors associated with them, are highly sensitive to definitions. Depression was the only symptom/life situation variable, and education the only sociodemographic measure, predictive across all three models. Risk indicators such as suicide thoughts and drinking problems were only significant for the traditional (mental health specialty) model. While patients within the traditional model definition were significantly younger than other enrollees, those within the model using the most expansive definition were significantly older. Varying the definition also led to different results in respect to experimental manipulations, geographic sites, and some specific types of comorbidity.

Conclusions. A reasonable definition, consistent with medical standards, requires, at least, a mental health diagnostic judgment and some form of psychotherapeutic or drug treatment. Studies of the content of mental health care are needed. (*Am J Public Health.* 1992;82:74–78)

Introduction

Many studies report a high prevalence of psychiatric disorder, particularly depression and anxiety, among patients treated in the general medical care sector.[1,2] One goal is to understand help-seeking patterns among persons with psychiatric need, to understand how physicians recognize psychiatric morbidity in their patients, and to determine how to better educate clinicians to appropriately identify and treat psychiatric disorder.[3,4] A large proportion of patients with depression in general medical care are undetected, probably more than half.[5] Depressive disorders, and even symptoms of depression short of the disorder threshold defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-III), are associated with considerable disability, comparable to or surpassing the disability associated with many serious medical conditions.[6]

In practice, data on the prevalence of psychiatric disorder in the general medical sector come from either small clinical studies of selected practice settings or large national data sets that are broadly representative but lack detail. The large data sets are a common source of estimates of psychiatric morbidity in medical contexts and of analyses of correlates of mental health use.

In major surveys such as the Epidemiological Catchment Area (ECA) study,[7] respondents having various psychiatric symptoms are asked whether they sought care for these symptoms and from whom. Information typically is not available on care actually provided or its intensity. For example, in the ECA study respondents were asked whether they talked about any problems they had with emotions or nerves that might have been connected to or in addition to the reason for their general medical care visit. This general response was then used as a basis for analyzing mental health use.

Analyses from the ECA study indicate that most people with a DSM-III diagnosis as assessed by the Diagnostic Interview Schedule have not sought care for their mental health problem in the prior 6 months, and among those who have, almost half received such care exclusively from the general medical care sector.[8,9] Persons with a mental health visit in the prior 6 months, relative to all respondents, were more likely to be young, White, female, and unmarried and to have had some college education. Such patients were also more, likely to use clinics, to have a regular source of medical care, and to be receptive to professionals.[10] An alternative source of data comes from log diaries completed by doctors, as in the National Ambulatory Medical Care Survey. In this survey the doctor is asked to record the patient's complaints, symptoms, or other reasons for the visit in the patient's own words, principal and other significant diagnoses, new or continued medications ordered or provided, nonmedication therapies such as psychotherapy and "other counseling," and the duration of the visit. Such data on instances of treatment are more detailed than those obtained in the ECA study but do not permit aggregation of visits by person.

The authors are with the Institute for Health, Health Care Policy, and Aging Research, New Brunswick, NJ.

Requests for reprints should be sent to David Mechanic, PhD, Institute for Health, Health Care Policy and Aging Research, Rutgers University, 30 College Avenue, New Brunswick, NJ 08903.

This paper was submitted to the journal July 11, 1990, and accepted with revisions May 6, 1991.

Jencks,[11] using data from the 1980 National Ambulatory Medical Care Survey weighted to reflect the US experience, found that "3.3% of visits [in primary] care were assigned a mental reason for visit, 5.6% received a mental diagnosis, 6.8% received a psychotropic drug, and 3.2% received psychotherapy/therapeutic listening" (p. 1904). Aggregating all of these indicators, excluding duplication, results in an estimate of 11.6% of visits during the year in which there is some evidence of mental health content. The concordance of these indicators, however, is much less common. Only one half of 1% were affirmative on all four criteria and 1.79% included a mental health reason, a mental diagnosis, and either the notation of drugs or psychotherapy. Thus using the inclusive rule of meeting any criterion yields a mental health visit rate 6.5 times greater than a definition based on the presence of a mental health reason, a diagnosis, and mental health treatment.

The conservative criterion is too limited since it is well established that patients with psychosocial distress may disguise the reason for their visit, presenting instead various physical complaints to justify the consultation.[12] Including visits that lead to both a diagnosis and some type of treatment essentially doubles the proportion of mental health visits from 1.79% to 3.49%, accounting for 29% of the mental health visits under the most lenient definition. Estimates of mental health visits are extremely sensitive to the classification assumptions used.

Conceptually, a mental health visit might be said to occur when a patient consults a clinician for emotional or psychophysiological symptoms. The clinician assesses the problem and provides care in direct response to the patient's complaints. Some patients, however, are confused about the sources of their distress, or may deny emotional symptoms. Thus, mental health visits also may include instances where a physician recognizes a psychiatric/emotional problem and provides treatment in response to this assessment, irrespective of the patient's definition, presentation, or understanding of the problem. It would ordinarily be expected that the doctor would make some diagnostic judgment before treatment, but the data indicate that some doctors treat without recording a diagnosis. Several credible alternative definitions are possible. The purpose of this analysis is to examine the extent to which there is stability in predictive models of use of mental health visits when definitions vary.

TABLE 1—Most Common Patterns of Mental Health Care for Total Sample 18 Years and Older in the RAND Health Insurance Experiment

n	Mental Health Specialty Provider	Mental Health Diagnosis	Psychotherapy	Psychotropic Drugs
288	No	No	No	Yes
267	Yes	Yes	Yes	No
219	No	Yes	No	Yes
216	No	Yes	No	No
172	Yes	Yes	Yes	Yes
45	Yes	Yes	No	No
36	Yes	No	No	No
25	No	Yes	Yes	No
21	Yes	Yes	No	Yes

Note. The total number of adults was 1324 (including patterns not shown).

Methods

Procedure

Our data source is the RAND Health Insurance Experiment (HIE), a large experimental study with random assignment carried out between 1974 and 1982. The HIE studied the impact of varying schedules of coinsurance on medical care use and other outcome variables.[13,14] Subjects were followed for 3 to 5 years, depending on experimental assignment, and six geographic sites were included in the study. Independent measures were collected at baseline from interviews and self-administered questionnaires and were linked with subsequent claims forms recording types of practitioners seen, the primary problem or diagnosis as well as other problems/diagnoses for which physicians supplied treatment, procedures used, and drugs prescribed. Measures of morbidity, based on International Classification of Diseases codes, were derived from the claims information. The classification of psychoactive drugs is based on work by the RAND researchers,[15] who classified antipsychotics, antidepressants, and minor tranquilizers, using the National Drug Code Directory, excluding prescriptions when the record showed it was in response to a physical problem. Psychiatric procedures were classified following the California Relative Value Codes. Since the vast majority of procedures constituted psychotherapy, we refer to them as psychotherapy here. Comorbidity measures from the claims forms, and variables characterizing the experimental groups and geographic sites used in the study, are included as controls. We test a model including risk indicators (suicide thoughts, legal problems, deviant sexual behavior, and drinking problems), measures of health status and life situation (depression, a general health index, and life events), comorbidity measures, and sociodemographic variables. The theoretical relevance of the model and the specific measures used are described elsewhere.[16] Our purpose here is to examine the stability of prediction under varying definitions, and, thus, we do not dwell on specific substantive findings. Since some respondents participated in the experiment for 3 years, and others 5 years, a variable (TERM) is included to correct for different periods of exposure.

The RAND researchers did extensive developmental work to ensure the reliability and validity of key questionnaire items, particularly those involving health perceptions.[17] Summaries of major results and significant publications from the HIE are available.[18,19]

Descriptive data are based on 4403 persons aged 18 to 61 at enrollment. In the multivariate analyses, data from Dayton, Ohio, are excluded because this site was used to pretest the instrument and many of the items in the survey are not comparable between Dayton and other sites. This reduced the sample to 3739, of which 3138 had complete information on all variables. The multivariate analysis is based on these 3138 respondents.

Descriptive Data and Study Definitions

Tables 1 and 2 show the most common patterns of mental health care for adults and children during the experiment. These data allow differentiating care provided by mental health professionals such as psychiatrists, psychologists, social workers, psychiatric nurses, or other mental health practitioners as compared with general medical personnel. Using any of four criteria for a mental health visit (seeing a mental health provider, receiving

TABLE 2—Most Common Patterns of Mental Health Care for Total Sample 5–17 Years of Age in the RAND Health Insurance Experiment

n	Mental Health Specialty Provider	Mental Health Diagnosis	Psychotherapy	Psychotropic Drugs
82	Yes	Yes	Yes	No
57	No	Yes	No	No
53	No	No	No	Yes
24	Yes	No	No	No
18	No	Yes	No	Yes
16	Yes	Yes	Yes	Yes
14	No	Yes	Yes	No
13	Yes	Yes	No	No

Note. The total number of children was 287 (including patterns not shown).

a mental health diagnosis, receiving psychotherapy, and receiving a psychotropic drug for conditions other than those justified by physical symptoms [e.g., back pain]), 1324 adults had a mental health visit over the period of the experiment (30.1% of the eligible sample). Only 42% of these adults had visits in the specialty mental health sector over the life of the experiment. Nine patterns of care account for 97% of all instances of care. Most common were visits to a non–mental health specialist who prescribed drugs either with or without a mental health diagnosis and visits to a mental health specialist who made a mental health diagnosis and provided psychotherapy but no drugs.

One hundred thirty-eight children, aged 5 to 17, received care in the specialty mental health sector, 48% of the 287 children who used mental health services. Eight patterns accounted for 97% of all mental health care. Most common patterns were visits to mental health specialists who made a diagnosis and provided psychotherapy, visits to generalists who made mental health diagnoses but provided no specific care, and visits to generalists who did not make a diagnosis but prescribed drugs. All but one of the children without a diagnosis but treated with drugs were seen exclusively by general physicians.

The first two tables illustrate that there are numerous ways of classifying and aggregating mental health visits, and the conventions in the literature vary depending on the opportunities allowable with existing data sources. Most would concede that the numbers of visits will vary widely depending on the definition used, but it is generally assumed that analytic work examining factors associated with such visits will not be affected substantially by definitional conventions or variabilities. Yet we know that the factors affecting help seeking in general are different from those that predict particular sources of care.[20,21] In the analysis that follows, we examine how varying definitions of a mental health visit affect conclusions about the correlates of care.

Using four criteria (provider type, diagnosis, psychotherapy, and psychotropic medications), three alternative definitions were derived from the claims files. The first definition includes all visits to mental health specialty providers that involve a mental health diagnosis and treatment with either drugs or psychotherapy, the traditional definition of a mental health visit. Four hundred sixty enrollees met these criteria over the course of the experiment (10% of the adult sample of 4403). A second definition includes enrollees who received a mental health diagnosis and who received either drugs or psychotherapy. This definition conceptually fits the characterization of the de facto mental health system. Seven hundred sixteen enrollees (16.3% of the total sample) met these criteria. The third definition—the most inclusive—includes the 1027 respondents who received either psychotherapy or a psychotropic drug regardless of the provider or an explicit mental health diagnosis (27.5% of the total sample). Generalists often treat psychologically distressed patients without making a diagnosis, and these criteria constitute a more robust definition of the de facto system.

Results

Table 3 presents three logistic regressions that compare individuals who made mental health visits as defined by our varying criteria (traditional, de facto, and robust de facto) with those who did not have a visit based on these criteria. The results vary greatly depending on the way a mental health visit is defined. In the comparison involving those who met the traditional criteria (Model 1), mental health visits are associated with depression, suicide thoughts, and drinking problems. Persons making such visits also have more education, are younger, and are less likely to be Black. None of the other sociodemographic variables are statistically significant. Mental health visits as defined in Model 1 occur significantly more frequently in some plans and geographic sites than others. Patients with comorbidity involving infections, nervous system disorders, digestive and respiratory problems, and accidents are significantly more likely to have such visits.

Model 2 presents results for those who had a mental health diagnosis and received a psychotropic drug or psychotherapy. Only depression is significant among the risk and health status/life situation variables. Most of the insurance plans and two of the sites are also statistically significant. Education is the only significant sociodemographic variable, and several comorbidity measures are significant, as well as the control for length of participation in the experiment (TERM).

Model 3, based on the most inclusive definition of mental health visits, includes all respondents in either sector who received either psychotherapy or a psychotropic drug. In this model a mental health diagnosis is not needed to meet criteria for a mental health visit. Depression is associated with a greater probability of having a visit, whereas the measure of favorable perceptions of general health status is associated with a lower probability. Older enrollees and those with more education are significantly more likely to have such visits. Persons making mental health visits by this definition are significantly more likely to have most comorbid conditions than those without such a visit.

Discussion

Definitions of mental health visits vary in the literature. Our analysis shows clearly that estimates of the number of mental health visits, and the factors associated with them, are highly sensitive to these definitions. Depression was the only symptom/life situation variable and education the only sociodemographic measure predictive in all three models. Several measures of comorbidity were consistent predictors, reinforcing the general awareness that comorbidity is common in mental health care. Several of the experimental variables were also significant across the three models.

More pertinent were the many important variables that had varying effects depending on definition. Risk indicators such as suicide thoughts and drinking problems were related only to the traditional specialty model, and this was the only model that had significantly fewer Black patients. Model 1 had significantly more younger patients, while Model 3 had significantly fewer younger patients. Enrollees with a lower perception of their general health status were significantly more common in Model 3, but not in the other models. The models also gave significantly different results in respect to some of the experimental manipulations, geographic sites, and some specific types of comorbidity.

The RAND HIE did not include all of the important measures found in other use studies, but there is substantial overlap. The discrepancy in results among models, and with other findings in the literature, argues for great care in defining precisely what we mean by mental health use. One important conclusion that emerges from this analysis is the critical importance of an appropriate definition of mental health use consistent with planning needs.

In a related paper, Mechanic[22] suggested that very broad definitions of mental health visits inadvertently encourage a perception that seriously overestimates the responsiveness of the medical care system to persons with mental health needs. These definitions, and the data that support them, take on a life of their own. The traditional definition, as defined in this paper, is probably too narrow in that it neglects the valuable services provided by generalists, but the robust de facto model is probably too lenient for a realistic assessment of service provision. Model 2 is, at least, a beginning consistent with medical standards. It is reasonable to expect that a physician will make a diagnostic judgment before initiating mental health treatment, and such an assessment is required under Model 2. Continued use of such data sets should be supplemented by qualitative studies that examine in detail the content and appropriateness of mental health care. □

TABLE 3—Logistic Regressions of Mental Health Visits on Risk Factors, Comorbidity, and Demographic Variables

Variable	Model 1 (specialist with diagnosis and drug or procedure) Coefficient	P	Model 2 (diagnosis with drug or procedure) Coefficient	P	Model 3 (drug or procedure) Coefficient	P
Health status/life situation						
Depression	.026	.000	.029	.000	.024	.000
General Health Index	.005	.267	−.001	.722	−.011	.002
Life events	−.085	.274	−.070	.300	.036	.533
Risk indicator						
Suicide thoughts	.342	.043	.092	.548	−.083	.566
Legal problems	.107	.335	.099	.332	−.007	.939
Deviant sexual behavior	.043	.538	.022	.739	.043	.467
Drinking problems	.336	.008	.218	.068	.051	.656
Sociodemographics						
Work	−.068	.638	−.074	.554	−.130	.251
Married	.176	.275	.180	.209	.077	.544
Black	−.833	.009	−.408	.067	−.310	.093
Male	.060	.710	−.152	.273	−.192	.116
Age	−.014	.040	.000	.954	.011	.035
Education	.010	.000	.050	.018	.045	.017
Income	−.106	.426	−.101	.384	−.000	.997
Comorbidity/term						
TERM	.119	.055	.109	.042	.053	.268
INFECT	.353	.008	.253	.030	.227	.032
MUSCLE	.224	.102	.402	.001	.550	.000
NERVOUS	.317	.035	.270	.040	.332	.004
DIGEST	.298	.043	.430	.001	.447	.000
SKIN	.243	.064	.312	.007	.286	.006
GENIT	.495	.001	.471	.000	.507	.000
RESPIT	.320	.018	.392	.001	.436	.000
ACCIDENT	.212	.118	.201	.086	.251	.015
ILLSYMP	.241	.083	.534	.000	.427	.000
CIRCUL	−.033	.827	.182	.155	.372	.001
NEOPLASM	−.006	.974	−.175	.240	.013	.923
Experimental manipulation or site						
Plan A	−.390	.101	−.568	.008	−.753	.000
Plan BD	−.825	.025	−.916	.004	−.979	.000
Plan EG	−.609	.230	−.876	.045	−.934	.013
Plan HJ	−1.03	.004	−.826	.004	−.797	.001
Plan KM	−1.09	.001	−1.097	.000	−.956	.000
Plan N	−.351	.169	−.491	.032	−.712	.001
Plan O	−.115	.566	−.249	.180	−.360	.037
Seattle, Wash	.390	.106	−.508	.016	−.429	.024
Frankburg, Mass	.025	.919	.149	.486	.072	.712
Charleston, SC	−.824	.016	−.641	.016	−.387	.090
Georgetown, SC	−.899	.005	−.476	.051	.025	.905
n	350		518		741	

Note. The total sample size was 3138.

Acknowledgments

This research was supported by National Institute of Mental Health Grant MH44654.

References

1. Morlock LL. Recognition and treatment of mental health problems in the general health care sector. In: Taube C, Mechanic D, Hohmann A, eds. *The Future of Mental Health Services Research.* Washington, DC: US Government Printing Office, 1989:39–61.
2. Regier DA, Goldberg ID, Taube CA. The de facto US mental health services system: a public health perspective. *Arch Gen Psychiatry.* 1978;35:685–693.
3. Goldberg D, Huxley P. *Mental Illness in the Community: The Pathway to Psychiatric Care.* New York, NY: Tavistock Publications; 1980.
4. Goldberg D. Reasons for misdiagnosis. In: Sartorius N, Goldberg D, de Girolamo G, et al., eds. *Psychological Disorders in General Medical Settings.* Toronto: Hogrefe and Huber; 1990:139–145.
5. Wells KB. *Depression as a Tracer Condition for the National Study of Medical Care Outcomes: Background Review.* Santa Monica, Calif: RAND Corp; 1985. R-3293-RWJ-HJK.
6. Wells KB, Stewart A, Hays RD, et al. The functioning and well-being of depressed patients: results from the medical outcomes study. *JAMA.* 1989;262:914–919.
7. Eaton WW, Kessler LG, eds. *Epidemiologic Field Methods in Psychiatry: The NIMH Epidemiologic Catchment Area Program.* New York, NY: Academic Press; 1985.
8. Shapiro S, Skinner EA, Kessler LG, et al. Utilization of health and mental health serv-

ices: three epidemiological catchment area sites. *Arch Gen Psychiatry.* 1984;41:971–978.
9. Shapiro S, Skinner EA, Kramer M, Steinwachs DM, Regier DA. Measuring need for mental health services in a general population. *Med Care.* 1985;23:1033–1043.
10. Leaf PJ, Livingston MM, Tischler GL, Weissman MM, Holzer CE, Myers JK. Contact with health professionals for the treatment of psychiatric and emotional problems. *Med Care.* 1985;23:1322–1337.
11. Jencks SF. Recognition of mental distress and diagnosis of mental disorder in primary care. *JAMA.* 1985;253:1903–1907.
12. Mechanic D. Patient behavior and the organization of medical care. In: Tancredi LR, ed. *Ethics of Health Care.* Washington, DC: Institute of Medicine, National Academy of Sciences; 1974.
13. Newhouse JP. A design for a health insurance experiment. *Inquiry.* 1974;11:5–27.
14. Lohr, KN, Brook RH, Kamberg, CJ, et al. Use of medical care in the RAND health insurance experiment: diagnosis- and service-specific analyses in a randomized controlled trial. *Med Care.* 1986;24(suppl):51–587
15. Wells KB, Manning WG Jr, Duan N, Ware JE Jr, Newhouse JP. *Cost Sharing and the Demand for Ambulatory Mental Health Services.* Santa Monica, Calif: RAND Corp; 1982. R-2960-HHS.
16. Mechanic D, Angel R, Davies L. Risk and selection processes between the general and the specialty mental health sectors. *J Health Soc Behav.* 1991;32:49–64.
17. Davies AR, Ware JE Jr. *Measuring Health Perceptions in the Health Insurance Experiment.* Santa Monica, Calif: RAND Corp; 1981. R-2711-HHS.
18. Manning WG, Newhouse JP, Duan N, et al. *Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment.* Santa Monica, Calif: RAND Corp; 1988. R-3476-HHS.
19. Newhouse J. Controlled experimentation as research policy. In: Ginzberg E, ed. *Health Services Research: Key to Health Policy.* Cambridge, Mass: Harvard University Press; 1991:161–194.
20. Greenley JR, Mechanic D. Social selection in seeking help for psychological problems. *J Health Soc Behav.* 1976;17:249–262.
21. Greenley JR, Mechanic D, Cleary PD. Seeking help for psychologic problems: a replication and extension. *Med Care.* 1987;25:1113–1128.
22. Mechanic D. Treating mental illness: generalist versus specialist. *Health Aff.* 1990;9:61–75.

Booklets Help Hospitalized Patients Quit Smoking

Quitting smoking may be the last thing on patients' minds when they enter the hospital. Yet as more and more medical institutions go smoke-free, patients must stop smoking when they enter the hospital.

Two booklets designed by the Hopkins Hospital Smoking Cessation Service help patients quit smoking, both during the hospital stay and afterward. The booklets are now available to health care providers elsewhere and can be personalized with a particular hospital's name on the cover.

As the titles suggest, *Quit Smoking for Good . . . While You're in the Hospital,* and *Stay Quit for Good . . . After You Leave the Hospital* offer patients tips on kicking the habit in both settings. Each booklet uses simple language, humor, and cartoon illustrations in a sympathetic approach to guiding patients through the quitting process. The booklets can be used as tools for nurses, hospital admissions officers, and addiction specialists.

A third booklet, *Help Someone You Care About Stay Quit for Good,* intended for family and friends of patients who help support the patient's smoke-free lifestyle will be available in January 1992.

The cost is $2.05 per booklet (discount available for large orders). To order call 301/955-4155.

Evaluation of a Self-help Dietary Intervention in a Primary Care Setting

Shirley A. A. Beresford, PhD, Elizabeth M. Z. Farmer, MA, Lisa Feingold, MSPH, Karen L. Graves, PhD, Sheron K. Sumner, PhD, and Richard M. Baker, MD

ABSTRACT

Background. Dietary intervention to reduce fat consumption and increase fiber consumption has been recommended by the National Cancer Institute, but there is little evidence concerning the effectiveness of self-help materials. The purpose of this study was to evaluate such self-help materials, introduced by a nurse in a primary care setting.

Methods. A randomized controlled trial involving 242 subjects was conducted in two primary care clinics in Chapel Hill, NC, in 1987. Changes in fat and fiber consumption in the intervention and control groups during the 3-month interval between interviews were compared using analysis of covariance.

Results. The estimated reduction in fat was 3.8g larger for the intervention group than for the control group, but the confidence interval included zero. For those individuals who had some responsibility for meal preparation there was a larger difference (−6.9g) in favor of the intervention group, although the difference using calorie-adjusted values was −3.8g with a 95% confidence interval (−7.1, −0.4). The differences for fiber change were smaller.

Conclusions. We found significant small but consistent differential changes associated with a minimal self-help intervention, but we cannot rule out the possibility of some response bias. Nonetheless, this study demonstrates that the use of self-help materials for dietary change is feasible, and may be effective. (*Am J Public Health.* 1992;82:79–84)

Introduction

Dietary modification to reduce fat consumption and increase fiber intake is one of the main preventive interventions recommended by the National Cancer Institute (NCI). The NCI's goal is to reduce breast and colon cancer mortality through the reduction of fat intake to 30% of total calories or less and through the increase of fiber consumption from 8 to 12 grams per day to 20 to 30 grams per day.[1]

Although these NCI recommendations have generally been well accepted, there have been few studies evaluating the efficacy of a particular intervention in modifying dietary fat and fiber intakes. The Ludwig Institute (Toronto) is evaluating the recurrence rates of polyps in patients who are counseled to follow a very high fiber, low fat diet compared with the rates for those who are counseled to follow Canadian Government dietary guidelines.[2] The University of Minnesota has evaluated different intervention strategies to modify fat and fiber, including correspondence courses, community interventions involving media,[3] and individual counseling using educational materials.[4] Results from the Women's Health Trial Vanguard Study indicate that intensive group participation was effective in maintaining a reduction of dietary fat intake over a 2-year period.[5]

A review of the cardiovascular disease prevention literature reveals that it is possible to influence the diets of individuals at higher risk for heart disease.[6] Four randomized studies published in the period 1975 to 1984 targeted subjects' saturated fat consumption in an attempt to reduce their risk of cardiovascular disease, and evaluated the impact of interventions on dietary fat consumption.[7–11] All four were successful in lowering fat consumption to some degree. The program with the least intensive intervention appeared to be self-teaching using a slide projector.[10] A more recent study in group general practice in Britain found individual counseling to be effective in the short term in lowering plasma lipid levels in men.[12]

No evidence has been published to date concerning the efficacy of self-help written materials, as an alternative form of self-teaching, in achieving behavior changes in dietary fat and fiber intake, although there is evidence for the usefulness of self-help materials in smoking cessation.[13–15] If parallels to the behavior changes associated with smoking cessation can be drawn, it appears that the influence of the primary care physician can be a useful adjunct to self-help materials.[16–19] Before researchers can enlist physicians to provide advice about di-

At the time of the study, Shirley A. A. Beresford was with the Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill. Elizabeth M. Z. Farmer and Lisa Feingold were with the Department of Epidemiology, School of Public Health, and Richard M. Baker was with the Department of Family Medicine, School of Medicine, all at the University of North Carolina at Chapel Hill. Karen L. Graves and Sheron K. Sumner were with the Department of Food-Nutrition-Food Service Management, School of Home Economics, at the University of North Carolina at Greensboro.

Requests for reprints should be sent to Shirley A. A. Beresford, PhD, Department of Epidemiology, SC-36, School of Public Health and Community Medicine, University of Washington, Seattle, WA 98195.

This paper was submitted to the journal July 28, 1990, and accepted with revisions May 7, 1991.

etary modification to their patients in conjunction with self-help materials, it is necessary to assess the feasibility and efficacy of this strategy. The present study was designed to meet this need.

Materials and Methods

Study Population and Data Collection

A randomized, controlled trial of patient self-help materials introduced by a nurse was conducted in two primary care clinics in Chapel Hill, NC, in 1987. The study population included adult patients (over 18 years) seeing their primary care physicians. Patients with life-threatening conditions or mental illness that would render them unable to respond to questions were excluded. Clinic days (Monday through Friday) were randomly assigned to intervention and control designations, and allocation was balanced in 10-day blocks. Three hundred forty patients waiting to see their physicians were invited by a research assistant to participate in the study, 167 on intervention days and 173 on control days. Twelve patients refused to participate. Patients who agreed to participate were interviewed in a small room, close to the waiting area. The interview questionnaire ascertained demographic information including educational level and whether the respondent had a home telephone. If the respondent had less than an eighth-grade education or had no home telephone, he or she was judged to be ineligible and the interview was terminated. Twenty percent of the intervention group and 23% of the control group were so excluded. For eligible respondents, the interview continued and information on current dietary habits was obtained. Some respondents were called to see their physicians before the interview was complete. Most of these (all but 14) were able to complete the interview and talk with the study nurse, if appropriate, after their appointments. The number of subjects successfully enrolled in the study was 242, 120 (72%) in the intervention group and 122 (71%) in the control group.

Those attending the clinic on an intervention day received the packet of self-help materials, which was briefly introduced to them by the study nurse. The subjects were then free to see their physicians as scheduled. Subjects seeing their physicians on control days received the baseline interview only, and then saw their physicians in the usual way. For subjects in the intervention group, use of the self-help materials at home was reinforced via a telephone call from the study nurse at about 10 days. Information concerning the acceptability and use of the materials was collected at this time. Approximately 3 months after the initial contact, both groups of subjects were telephoned by the research assistant, who conducted a follow-up telephone interview that included a repeat assessment of current fat and fiber consumption and some new questions concerning food choices. The research assistant was not informed as to the group to which an individual subject belonged.

Intervention

The intervention in this study was self-help materials, introduced to the subject by a study nurse and reinforced later by telephone. This registered nurse was not a member of the health care team. The presentation by the nurse was limited to 5 minutes, to resemble the kind of time commitment feasible for primary care physicians. The self-help materials were developed by some of us (K. G., S. S.), revised to an eighth-grade reading level, and professionally printed in attractive folders with 15 card inserts of varying sizes to give the user easy access to different types of information. The self-help materials emphasized a healthy diet for all individuals, and focused on fat and fiber. The same information was presented in several different ways in the packet. Included in the packet were (1) a checklist to help the individual determine the extent to which he or she had made desirable changes in fat and fiber consumption; (2) a sample food record with a fat and fiber counter booklet to help determine exactly how much fat and fiber had been consumed; and (3) menus offering a variety of food choices to satisfy personal tastes while modifying fat and fiber intake to recommended levels. Information about the NCI guidelines was given.

Measurement of Outcome

The outcome of interest in this study is patients' dietary behavior with respect to fat and fiber intake approximately 3 months after an index visit to their physicians. The accurate measurement of dietary behavior has been a problem of concern to nutritionists and epidemiologists for a long time. Methods that can accurately classify an individual's dietary intake require a substantial commitment from either the interviewer or the individual respondent. For population studies linking dietary factors to disease, less accurate methods, which impose fewer demands on the interviewer and respondent, are acceptable.[20] These methods include the Semiquantitative Food Frequency Questionnaire (SQFFQ) developed by Willett,[21] which was used in the present study. This questionnaire has been validated against 4-day food records, a method that has greater face validity, both in a small group of healthy volunteers[22] and in 173 nurses.[23] Little information is available about the sensitivity of this food frequency method in detecting changes in diet that might be expected as the result of a dietary intervention. As with some physiological measures, such as systolic blood pressure, the within-person variability in dietary behavior may be at least as great as the between-person variability.[23] With this in mind, two classes of outcome measures were chosen in this study.

The first was fat and fiber consumption from the SQFFQ,[21] which was administered during an interview both at baseline and at follow-up. The values for fat, fiber, and calories for each food item or class of food items were estimated by one of us (S.S.), using the most recent analyses from USDA Handbook 8.[24] The analyses for the main study (described below) were then reworked, using the database and modifications developed by Willett[21] for fat and calories, and these analyses are presented here. There was a very high correlation between the two methods, and the results did not change substantially. For dietary fiber, the most recent data from Bowes and Church were used,[25] rather than Willett's database.

The second outcome measure was fat and fiber scores from a behavioral questionnaire designed to detect changes in food choices (as opposed to relative consumption in a population), which was administered at follow-up only. This behavioral questionnaire was developed from the checklist used in the self-help materials.

Validation Studies

Two validation studies were conducted to support the use of the instruments in this study. The first study examined the concordance of results obtained face to face and those obtained by telephone. Thirty-two ward clerks at North Carolina Memorial Hospital were asked to complete the SQFFQ on two occasions, 1 week apart. Complete information was obtained from 30. One occasion was an in-person interview (to mimic the baseline interview in the main study) and the other was a telephone interview. A crossover design was used to randomly allocate par-

ticipants to order of administration. Effects were computed by analysis of variance for split-plots.[26] Interactions between mode and period were assessed for both fat and fiber intake. There was no interaction between period and mode for either daily fat or daily fiber intake, and there was no significant difference due to mode. Indeed, the correlation between fat intake as assessed by telephone and fat intake as assessed face to face was of the same order as other test-retest correlations using the same mode of administration on two occasions, namely, 0.5.[22] On the other hand, there was a significant period effect, with responses to the first questionnaire indicating higher intake for both fat and fiber, regardless of mode of administration. This effect has been found in other studies.[27,28]

In a second validation study, the behavioral measure was compared with a 3-day food record (a near–gold standard measure of dietary intake) among monitor and division clerks at Moses H. Cone Memorial Hospital. Scatter plots and correlation coefficients were used to compare fat and fiber scores from the two measures. Twenty-nine ward clerks volunteered to participate. Summing the desirable responses ("always" or "frequently" for some questions; "rarely" or "never" for others) into a low-fat-behavior score and a high-fiber-behavior score resulted in scores ranging from 0 to 12 and 0 to 6, respectively. The low-fat-behavior score was moderately negatively correlated (as expected) with daily fat intake as assessed by both the SQFFQ (−0.44) and the 3-day food record (−0.31). The corresponding correlations between the high-fiber-behavior score and daily fiber intake were 0.47 and 0.45, respectively.

Statistical Methods (Main Study)

For the main study, the intervention and control groups' changes in fat and fiber consumption in the 3-month interval were compared. Descriptive analyses were conducted for daily fat and fiber consumption at baseline and at follow-up by examining the separate contributions to the daily total from different food groups. Responses to individual food-choice questions on the behavioral measure were also examined.

Analysis of covariance was used, with the follow-up value as the dependent variable and the baseline value as a covariate. Fat consumption was assessed both in grams and adjusted for calories consumed. The adjustment method[29] used simple linear regression coefficients of fat on calories at baseline for the baseline value and fat on calories at follow-up for the follow-up value. The simple regressions used data on all 189 subjects with complete information for caloric intake. A similar analysis was carried out for dietary fiber. Additional covariates were potential confounders such as age, education, gender, race, who prepared meals, household size, number of cigarettes smoked, miles from the clinic, and clinic, which were assessed for confounding and successively eliminated backward from the model. One non-Black non-White subject was excluded from these analyses. Variables that were significantly associated with the dependent variable or whose removal from the model caused a change of more than 10% in the partial regression coefficient for the group variable were retained in the model. The potential confounders age, education, gender, race, and who prepared meals were assessed as potential effect modifiers first, by including their interaction with group in the full model. All such two-way interactions were assessed together, using a chunk test.[30] Confounding variables in the final model were reassessed individually as effect modifiers by adding a two-way interaction term with group. A consistent set of confounders or potential confounders was sought for the different approaches to the analyses for fat and fiber. From the final models, the estimated group differences for change (follow-up minus baseline) were calculated for purposes of presentation, together with their 95% confidence intervals. Analyses were repeated, restricting observations to those whose follow-up interview was completed within 5 months of the baseline interview. Essentially the same results were obtained.

Similar analyses were conducted using total scores of desirable responses to the fat food-choice questions and the fiber food-choice questions that were part of the behavioral measure at follow-up. The behavioral measure was not used at baseline, so no adjustment could be made for baseline values. The validity of these analyses relies on the comparability of the two groups at baseline.

TABLE 1—Baseline Information by Group for Those Who Completed Follow-up

	Intervention (n = 93)	Control (n = 98)
Age, y	42	44
% Black	35	32
% Educated < 12 y	43	51
Fat (g)	66	67[a]
Fiber (g)	14	15
Fat/calories (%)	32	31[a]

[a]Based on one fewer observation.

Results

In the main study, the two groups of subjects appeared similar with respect to demographic and dietary intake information at baseline, and 78% of the intervention group and 80% of the control group completed the follow-up telephone interview. Those lost to follow-up in both groups included a higher proportion of Black respondents and those with less than a 12th-grade education. The baseline characteristics of those who completed follow-up in each group are shown in Table 1.

The baseline percentage of calories from fat is surprisingly low in both groups. Such an underestimation of fat calories was not found in the Moses H. Cone Memorial Hospital validation study, where the percentage of calories from fat was estimated at 39%, using data from either the food record or the SQFFQ. The participants in the Moses H. Cone study reported a higher frequency of consumption of baked goods than did participants in the main study, which could account for this discrepancy. In any event, the underestimates, if so they were, of fat calorie consumption in the main study applied equally to the intervention and control groups.

One hundred and ninety-one subjects, 93 in the intervention group and 98 in the control group, were followed up by telephone approximately 3 months after the baseline visit. The daily fat intake decreased between baseline and follow-up in both groups, and this decrease was accompanied by a decrease in kilocalories in both groups. Daily fiber intake in the control group decreased, on average, over the

TABLE 2—Proportions with Desirable Response to Questions about Food Choices at Follow-up

"How often do you choose . . ."	Desirable Response	Intervention Group	Control Group	Estimated Difference (%)	95% CI
. . . breakfast cereals such as bran cereals, oatmeal, or other whole grain choices?	Always or Frequently	59/91	55/95	7	(−7%, 21%)
. . . breads or crackers made from whole grains (such as wheat, rye, or bran)?	Always or Frequently	65/91	56/98	14	(1%, 28%)
When you eat fruits and vegetables . . . ones with edible skins, peels, or seeds?	Always or Frequently	65/93	59/98	10	(−4%, 23%)
. . . to eat less than five servings of fruits and vegetables per day	Rarely or Never	30/93	25/98	7	(−6%, 20%)
. . . low-fat milk (2% or skim) for cooking and drinking?	Always or Frequently	55/92	54/94	2	(−7%, 21%)
When you eat meat . . . to trim the visible fat?	Always or Frequently	74/91	70/96	8	(−4%, 20%)
When you eat chicken or other poultry . . . to eat it with the skin on?	Rarely or Never	44/89	42/94	5	(−6%, 23%)
When meat, poultry, or fish are cooked in your home . . . to broil, stir-fry, or bake?	Always or Frequently	64/91	50/96	18	(5%, 32%)

TABLE 3—Adjusted Differences between Groups (Intervention − Control)[a]

	Group Difference	95% CI
Change in daily fat	−3.8 g	−9.4, 1.9
Change in calorie-adjusted fat	−2.0 g	−4.9, 0.9
Change in daily fiber	0.6 g	−0.9, 2.1
Change in calorie-adjusted fiber	0.9 g	−0.4, 2.1

Note. n = 190 randomized patients available for follow-up. Four individuals had missing values for covariates; one had missing values for fat and calories.
[a]Differences were adjusted for baseline value, miles from clinic, who prepared meals, and number of cigarettes smoked.

period of follow-up, while fiber intake increased in the intervention group.

Food choices, ascertained from the behavioral questionnaire at follow-up, are presented in Table 2. For all questions, the percentage with a desirable response was consistently greater in the intervention than in the control group. The greatest differences in behavior changes were observed in choice of cooking methods and choice of whole-grain breads or crackers. These changes were reflected in analyses of covariance examining number of low-fat behaviors (a maximum of 12) and high-fiber behaviors (a maximum of 6), adjusting for race, who prepared meals, and number of cigarettes smoked. The estimated group difference in low-fat–behaviors was 0.3 (95% confidence interval −0.4, 1.1); the difference in high-fiber–behaviors was 0.6 (95% confidence interval 0.1, 1.0).

The results of the main analyses of covariance are presented in Table 3. The estimated change in fat consumption in grams was 3.8 g larger for the intervention group than for the control group, adjusted for the four covariates, but the confidence interval included both positive and negative values. Generally, the results for both fat and fiber are not large, although they are consistent with a small change in favor of the intervention group.

There was some indication of possible effect modification with who prepared meals, so separate analyses were conducted for the subgroup that reported no responsibility for meal preparation and for the subgroup who reported shared or total responsibility. Those with no responsibility for meal preparation were a small subgroup (n = 43); the estimated group differences were small and the confidence intervals were wide.

On the other hand, for the larger subgroup of those with at least some responsibility for meal preparation, the estimated group difference after adjusting for baseline value and other covariates was −6.9 g, which was reduced to a difference of −3.8 g when calorie-adjusted values were used. The associated 95% confidence intervals are −13.2, −0.6 and −7.1, −0.4, as shown in Table 4. These estimates represent the greater decrease in fat intake in the intervention group beyond that in the control group over the period of follow-up.

Discussion

We found small but consistent differential changes in consumption of both dietary fat and dietary fiber over a 3-month period associated with a minimal self-help intervention. The differential change in fat consumption was statistically significant in the subgroup of respondents with full or shared responsibility for meal preparation. Indeed, it might have been expected that this subgroup would experience a larger differential change than those with no responsibility for meal preparation. The degree of differential change was of the same order of magnitude as that found among the husbands of the participants in the Women's Health Trial Vanguard Study,[31] namely, about 2% calories from fat. It is not surprising that this change is less than that found among the participants (the women) in the Women's Health Trial, since that intervention was rather intensive, consisting as it did of regular group sessions meeting over a period of 2 years.[5]

The participants in our study were not selected on the basis of risk for disease

or because they had a particular level of dietary fat or fiber consumption. Although potential subjects who had less than an eighth-grade education were excluded (because our intervention was based on written materials), a wide range of educational levels were included in the study. Those who were followed throughout the study were not substantially different from those who enrolled at baseline. The crossover validation study demonstrated that results from telephone administration of the SQFFQ are comparable to results from face-to-face administration with respect to fat and fiber consumption.

One of the limitations of the SQFFQ is that individuals tend to overestimate their consumption of a number of foods that are specified on the questionnaire. The instrument itself may lead to underestimation of nutrients because of the limited number of foods (152 items in 108 groups of foods) included on the questionnaire. It may be reasonable to assume that such overestimation and underestimation would be similar from one occasion to another, or from one group to another. Results from the validation study of ward clerks at North Carolina Memorial Hospital call into question the first of these assumptions, since it appears that the estimates of consumption were greater on initial assessment than 1 week later. The second assumption may also be questioned if, at follow-up, members of the intervention group are more aware of the desirable responses than are members of the control group, thus creating a response bias. The analysis of grams of dietary fat, adjusting for baseline responses, makes some adjustment for overreporting, but the analysis of calorie-adjusted values adjusts for general over- or underreporting directly at both follow-up and baseline. Arguably, greater weight should be placed on the estimates obtained using these calorie-adjusted values.

Because of the difficulty of measuring dietary behaviors of groups with sufficient accuracy that small changes in behavior can be detected, we used a number of different measures in this study. The consistency of the changes in favor of the intervention group, particularly with respect to dietary fat reduction, is striking. On the other hand, the leap of inference from finding changes associated with the intervention to attributing them to the intervention and concluding that it was effective in changing diet has to be made with caution. None of the outcome measures is free from potential response bias, and the measure of food choices, where evidence supporting differential changes in fiber consumption is strongest, is particularly liable to bias. The wording of the food-choice questions is fairly similar to that on one of the self-monitoring sheets in the package of self-help materials. Fabrication of responses to the SQFFQ would be harder.

Nonetheless, by using the public health approach of attempting to shift fat consumption down and fiber consumption up by small amounts in the whole population, we have demonstrated that the use of self-help materials for dietary change is feasible and may be effective. For such materials to be useful in primary care practice, they must be introduced to the patient by a member of the practice team during the normal course of an office visit; therefore, there is a need for a further study in which the influence of the physician or nurse or both is added to the intervention. Such a study is now in progress in western Washington. The results from this study should provide evidence about the benefits of a rather simple intervention that could easily be incorporated into the normal practice of a primary care clinic. □

TABLE 4—Adjusted Differences between Groups (Intervention – Control)[a] among those who had some responsibility for preparation of meals.

	Group Difference	95% CI
Change in daily fat	−6.9 g	(−13.2, −0.6)
Change in calorie-adjusted fat	−3.8 g	(−7.1, −0.4)
Change in daily fiber	0.7 g	(−1.0, 2.4)
Change in calorie-adjusted fiber	1.3 g	(−0.1, 2.7)

Note. n = 146 randomized patients available for follow-up who had some responsibility for preparation of meals; three individuals had missing values for covariates; one had missing values for fat and calories.
[a]Differences were adjusted for baseline value, miles from clinic, and who prepared meals.

Acknowledgments

This work was funded in part by the Centers for Disease Control's Prevention Centers Grant to the University of North Carolina (UNC). Portions of this manuscript were presented at the International Epidemiological Association meeting, Los Angeles, CA, August, 1990. We gratefully acknowledge the support of Dr Suzanne W. Fletcher and Dr Barbara S. Hulka during the development of this project, and of Dr O. Dale Williams as Director of the UNC Center for Health Promotion and Disease Prevention. We are grateful to Kathleen Welshimer, Kathleen Girardin, Sharon Campbell, Michelle Carnaghi, and Carolyn Dunn for their assistance with various parts of this project. Finally, we thank the patients and ward clerks of North Carolina Memorial Hospital and the ward clerks of Moses Cone Memorial Hospital, without whose participation the study would not have been possible.

References

1. *Cancer control objectives for the nation: 1985–2000.* Bethesda, Md: National Cancer Institute, 1986;2:1–105. NCI Monographs.
2. McKeown-Eyssen GE. Fiber intake in different populations and colon cancer risk. *Prev Med.* 1987;16:532–539.
3. Potter J, Graves K, Finnegan J, et al. The Cancer and Dietary Intervention Project: a community-based intervention to reduce nutrition-related cancer risk. *Health Educ Res.* 1990;5:489–503.
4. Allen SS, Froberg D, McCarthy P, Myers S, Hunninghake DB. Preferences and opinions of consumers vs dietitians on cholesterol education materials. *J Am Diet Assoc.* 1991;91:604–606.
5. Henderson M, Kushi L, Thompson D, et al. Feasibility of a randomized trial of a low-fat diet for the prevention of breast cancer: dietary compliance in the Women's Health Trial Vanguard Study. *Prev Med.* 1990;19:115–133.
6. Baron J, Gleason R, Crowe B, Mann J. Preliminary trial of the effect of general practice based nutritional advice. *Br J Gen Pract.* 1990;40:137–141.
7. Bender DE, Beresford SAA, McFall SL. Can dietary behavior be changed? A review of interventions aimed at controlling hypertension. *J Comm Health.* 1989;14:18–43.
8. Hjermann I. Intervention on smoking and eating habits in healthy men carrying high risk for coronary heart disease: The Oslo Study. *Acta Med Scand.* 1981;651(suppl):281–284.
9. Hjermann I, Velve Byre K, Holme I, Heren P. Effect of diet and smoking intervention on the incidence of coronary heart disease. *Lancet.* 1981;ii:1303–1310.
10. Neaton JD, Broste S, Cohen L, Fishman EL, Kjelsberg MO, Schoenberger J. The multiple risk factor intervention trial (MRFIT), VII: a comparison of risk factor changes between the two study groups. *Prev Med.* 1981;10:519–543.
11. Mojonnier MH, Hall Y, Berkson DM, et al. Experience in changing food habits of hyperlipidemic men and women. *J Am Diet Assoc.* 1980;77:140–150.
12. Menotti A. An Italian preventive trial of coronary heart disease: the Rome project

of coronary heart disease prevention. *Prog Biochem Pharmacol.* 1983;19:230–244.
13. Glasgow RE. Effects of self-control manual, rapid smoking and amount of therapist contact in smoking cessation programs. *J Consult Clin Psychol.* 1978;46:1439–1447.
14. Glasgow RE, Schafer L, O'Neill H. Self-help books and amount of therapist contact in smoking cessation programs. *J Consult Clin Psychol.* 1981;49:659–667.
15. Schneider SJ, Benya A, Saiger H. Computerized direct mail to treat smokers who avoid treatment. *Comput Biomed Res.* 1984;17:409–418.
16. Pederson LL. Compliance with physician advice to quit smoking: a review of the literature. *Prev Med.* 1982;11:71–84.
17. Ockene JK. Physician–delivered intervention for smoking cessation: strategies for increasing effectiveness. *Prev Med.* 1987;16:723–737.
18. Janz NK, Becker MH, Kirscht JP, Eraker SE, Billi JE, Woolliscroft JO. Evaluation of a minimal-contact smoking cessation intervention in an outpatient setting. *Am J Public Health.* 1987;77:805–809.
19. Thompson RS, Michnich ME, Friedlander L, Gilson B, Grothaus LC, Storer B. Effectiveness of smoking cessation interventions integrated into primary care practice. *Med Care.* 1988;26:62–76.
20. Block G. A review of validations of dietary assessment methods. *Am J Epidemiol.* 1982;115:492–505.
21. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semi-quantitative food frequency questionnaire. *Am J Epidemiol.* 1985;122:51–65.
22. Willett WC, Reynolds RD, Cottrell-Hoehner S, Sampson L, Browne ML. Validation of a semi-quantitative food frequency questionnaire: comparison with a 1-year diet record. *J Amer Diet Assoc.* 1989;87:43–47.
23. Sempos CT, Johnson NE, Smith EL, Gilligan C. Effects of intraindividual and interindividual variation in repeated dietary records. *Am J Epidemiol.* 1985;121:120–130.
24. *Composition of Foods—Raw, Processed and Prepared.* Washington, DC: US Dept of Agriculture; 1963. USDA Handbook 8.
25. Pennington J, Church H. *Bowes and Church's Food Values of Portions Commonly Used.* 15th ed. Philadelphia, PA: JB Lippincott; 1989.
26. Hills M, Armitage P. The two-period cross-over clinical trial. *Brit J Clin Pharmac.* 1979;8:7–20.
27. Bloemberg BPM, Kromhout D, Obermann-De Boer GL, Van Kampen-Donker M. The reproducibility of dietary intake data assessed with the cross-check dietary history method. *Am J Epidemiol.* 1989;130:1047–1056.
28. Hankin JH, Yoshizawa CN, Kolonel LN. Reproducibility of a dietary history in older men in Hawaii. *Nutr Cancer.* 1990;13:129–140.
29. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. *Am J Epidemiol.* 1986;124:17–27.
30. Kleinbaum DG, Kupper LL. *Applied Regression Analysis and Other Multivariable Methods.* Boston, Mass: Duxberry Press; 1978.
31. White E, Hurlich M, Thompson RS, et al. Dietary changes among husbands of participants in a low-fat dietary intervention. *Am J Prev Med.* 1991;7:319–325.

American Journal of Epidemiology Corrects Its Press Release

"New Insights into Sudden Infant Death Syndrome," a news release published on p 1179 of the September 1991 issue of the *American Journal of Public Health* (Vol 81, No. 9) included an inaccurate statement.

Paragraph 5, sentence 2, as printed, reads: " 'The more common form of breech delivery called 'frank' presentation, with the baby emerging buttocks first, was not associated with an increased risk for SIDS,' Buck says." The corrected sentence reads as follows: " 'The sole control breech delivery was the more common delivery called 'frank' presentation, with the baby emerging buttocks first,' Buck says."

The original article, of which Germaine M. Buck, PhD, is the lead author, was published in the May 1, 1991, issue of the *American Journal of Epidemiology*. For questions, please contact the author at the Department of Social and Preventive Medicine, School of Medicine, SUNY at Buffalo.

Public Health Briefs

Cigarette, Alcohol, and Coffee Consumption and Spontaneous Abortion

Ben G. Armstrong, PhD, Alison D. McDonald, MD, and Margaret Sloan, BA

ABSTRACT

We analyzed data from a survey of occupational factors and pregnancy outcome to examine the effects of cigarette, alcohol, and coffee consumption on pregnancy outcome. Clear and statistically significant associations were found between cigarette and alcohol consumption and spontaneous abortion. There was a weaker but statistically significant association with coffee consumption: If the associations were casual, 11% of the spontaneous abortions could be attributed to smoking, 5% to alcohol, and 2% to coffee. (*Am J Public Health.* 1992;82:85–87)

Introduction

An association of spontaneous abortion has been reported with smoking[1–3] and with alcohol,[3,4] but not with coffee. To study these relationships, we further analyzed data collected for a large study of occupational factors and pregnancy outcome.[5]

Methods

Some 56 000 women were interviewed, comprising all who had had a delivery or a spontaneous abortion in 11 Montreal hospitals in a 2-year period (1982 to 1984). Questions on occupational, personal, and social factors were asked for all recently completed (current) and all previous pregnancies. Details of the study design and methods have been fully described elsewhere.[6,7]

Among current pregnancies, the spontaneous abortions occurred to women treated in hospital, and thus were only a proportion of the total and not representative. Previous pregnancies, in which spontaneous abortions are more fully reported,[8] were used in these analyses. For each pregnancy, an inquiry was made about coffee, alcohol, and cigarette consumption during the first trimester. The numbers of cigarettes and cups of coffee per day and drinks of each type per week were recorded. In calculating total alcohol consumption, one and a half glasses of wine were taken as equivalent to a glass of beer or measure of spirits ("one drink").

All previous pregnancies were included, regardless of whether the woman was employed. Abortions induced for personal reasons or for fetal defect were excluded. Of a total of 48 582 previous pregnancies, 1436 were excluded because of missing information, leaving 47 146 for analysis. However, as questions on coffee consumption during previous pregnancies were introduced after the survey began, information on this factor was available for only 35 848 pregnancies.

Possible confounding by maternal age (at each pregnancy outcome), educational level, ethnic group, and employment during pregnancy was controlled by including these factors together with those under study in logistic regression analyses of spontaneous abortion risk. Allowance was made for dependence between the outcomes of successive pregnancies in the same woman by including the number of previous live births and spontaneous abortions as a risk factor.[9] Risks attributable to specific factors were estimated from logistic regression coefficients.[10]

Results

The overall rate of spontaneous abortion in previous pregnancies was 21.6%; it varied with the factors identified as potential confounders to the extent shown in Table 1.

After control for these factors, there remained a strong association of smoking habit with abortion risk (see Table 2), compatible with a linear trend on the logistic scale in which odds ratios increased by a factor of 1.20 (1.18–1.23) for each 10 cigarettes per day.

The authors are with the School of Occupational Health, McGill University, Montreal, Quebec.

Requests for reprints should be sent to Ben G. Armstrong, PhD, School of Occupational Health, McGill University, 1130 Pine Avenue West, Montreal, Quebec, Canada H3A 1A3.

This paper was submitted to the journal July 19, 1990, and accepted with revisions May 6, 1991.

TABLE 1—Rates and Odds Ratios for Spontaneous Abortion (SA) by Maternal Characteristics

	No. Pregnancies	No. SAs	%	OR	95% CI
Age					
< 20 years	4 802	947	19.7	1.00	
20–29	35 500	7 057	19.9	0.88	0.81–0.96
30–34	5 774	1 713	29.7	1.37	1.24–1.52
35+	1 070	474	44.3	2.68	2.31–3.11
Number and outcome of previous pregnancies					
0	29 890	5 321	17.8	1.00	
1, 0 SA	8 705	2 169	24.9	1.56	1.47–1.65
1, 1 SA	2 855	826	28.9	1.82	1.67–1.99
2, 0 SA	2 098	596	28.4	1.81	1.63–2.01
2, 1 SA	1 590	564	35.5	2.44	2.19–2.73
2, 2 SA	456	207	45.4	3.69	3.05–4.46
3+, 0 SA	564	112	19.9	1.14	0.92–1.42
3+, 1 SA	516	193	37.4	2.59	2.15–3.12
3+, 2 SA	340	142	41.8	3.05	2.44–3.81
3+, 3+ SA	132	61	46.2	3.82	2.69–5.42
Ethnic group (color/language)					
White/English	8 959	1 826	20.4	1.00	
White/French	26 951	6 016	22.3	1.14	1.07–1.21
Other	11 236	2 349	20.9	1.06	0.99–1.14
Education completed					
< Secondary	13 619	2 873	21.1	1.00	
< College	18 977	4 056	21.4	1.10	1.04–1.17
College	14 550	3 262	22.4	1.14	1.07–1.22
Employment Status					
Unemployed	20 377	4 421	21.7	1.00	
Employed	26 769	5 770	21.6	1.11	1.05–1.16
All	47 146	10 191	21.6		

Note. OR = odds ratio, estimated by logistic regression with all factors, including cigarette, alcohol, and coffee consumption, in the model. CI = confidence interval.

TABLE 2—Rates and Odds Ratios for Spontaneous Abortion by Cigarette, Alcohol, and Coffee Consumption

	No. Pregnancies	Spontaneous Abortions			
		n	%	OR	95% CI
Cigarettes (per day)					
Nonsmoker	31 813	6 364	20.0	1.00	
1–9	2 861	572	20.0	1.07	(0.97–1.18)
10–19	4 987	1 100	22.1	1.22	(1.13–1.32)
20+	7 485	2 155	28.8	1.68	(1.57–1.79)
Alcohol (drinks per week)					
None	33 164	6 793	20.5	1.00	
1–2	9 099	2 068	22.7	1.11	(1.05–1.18)
3–6	3 069	776	25.3	1.23	(1.13–1.34)
7–20	1 527	456	29.9	1.47	(1.31–1.65)
21+	287	98	34.1	1.82	(1.21–2.34)
Coffee (cups per day)					
None	14 672	2 987	20.4	1.00	
1–2	15 035	3 196	21.3	0.98	(0.93–1.04)
3–4	4 068	981	24.1	1.02	(0.94–1.12)
5–9	1 574	442	28.1	1.17	(1.03–1.32)
10+	499	154	30.9	1.19	(0.97–1.45)
All	47 146	10 191	21.6		

Note. OR = odds ratio, estimated by logistic regression with all factors, including age, pregnancy history, ethnic group, education, and employment status, in the model. CI = confidence interval.

The consumption of alcohol was also clearly associated with elevated risk; odds ratios increased on average by a factor of 1.26 (1.19–1.33) for each drink per day. The trend in risk was not smooth but rose abruptly between 0 and 1 to 2 drinks per week, leading to a significant departure from linearity ($P = .03$).

The risk by coffee consumption increased on average by a factor of 1.017 (1.004–1.030) for each cup per day; this trend was also statistically significant ($P = .01$).

Modification of these three effects by other variables (interaction) was systematically investigated. There were a few statistically significant interactions (smoking–age, smoking–alcohol, and coffee–gravidity), but their magnitude was small. Information on these interactions is available from the authors on request.

Discussion

The overall spontaneous abortion rate (21.6%) was high, even compared with that in other studies based on mothers' recall, in which the rates were around 15%.[11–13] This high rate may be explained by women's being more likely to plan further pregnancies after a spontaneous abortion than after a live birth; pregnancies preceding a current pregnancy will thus overrepresent spontaneous abortions.

The associations observed may have been affected by biases, despite the allowance made for many confounders. The retrospective nature of the enquiry could have led to biased reporting of the use of alcohol, cigarettes, or coffee, or recall of spontaneous abortion. The plausibility of such bias is somewhat controversial,[14] but the fact that these habits are not generally believed to be associated with abortion makes bias less likely. However, the abortions occurred somewhat closer in time to the interviews than did live births, so if recall was time-related, some bias might result.

The association observed between abortion and cigarette smoking is similar to that found in other studies, with magnitude of risk in the middle of the reported range.[1–3]

The risk associated with the moderate consumption of alcohol in the present study (a factor of 1.26 per daily drink) was less than those found in other studies; for example, one study found that risk increased by a factor of 2.33 for women drinking between two and six drinks per week,[4] and another found risk increased by a factor of 1.98 for one to two drinks per day.[3] Previous studies were either case-control with hospitalized abortions as cases,[4] or prospective, following women from their first antenatal visit.[3] Differ-

ences might thus be related to biases specific to the study designs.

The small risk associated with coffee consumption might be explained by residual confounding. If caffeine were a causal agent, we would have suffered loss of power from lack of information on other caffeine intake. However, nutrition survey data show total caffeine intake and coffee consumption to be highly correlated ($r = .98$) in Quebec women (Health and Welfare Canada, personal communication). Evidence of an association between caffeine consumption and early fetal loss has been reported,[15] but no such association has been reported, so far as we are aware, with spontaneous abortion.

If the observed associations were causal, cigarettes accounted for about 11% of all spontaneous abortions (40% in women who smoked 20 or more cigarettes per day), alcohol consumption for about 5% (45% in women drinking 3 or more drinks per day) and coffee for about 2% (16% in women drinking 10 cups per day). These risks warrant concern among public health professionals.

References

1. Kline J, Stein Z, Susser M, Warburton D. Smoking: a risk factor for spontaneous abortion. *N Engl J Med.* 1977;297:793–796.
2. Himmelberger D, Brown B, Cohen E. Cigarette smoking during pregnancy and the occurrence of spontaneous abortion and congenital abnormality. *Am J Epidemiol.* 1978;108:470–479.
3. Harlap S, Shiono PH. Alcohol, smoking, and incidence of spontaneous abortions in the first and second trimester. *Lancet.* 1980;2:173–176.
4. Kline J, Shrout P, Stein Z, Susser M, Warburton D. Drinking during pregnancy and spontaneous abortion. *Lancet.* 1980;2:176–180.
5. McDonald AD. Work and pregnancy. *Br J Ind Med.* 1988;45:577-580. Editorial.
6. McDonald A, McDonald J, Armstrong B, Cherry NM, Delorme C, Robert D. Occupation and outcome of pregnancy. *Br J Ind Med.* 1987;44:521–526.
7. McDonald AD, McDonald JC, Armstrong BG, et al. Fetal death and work in pregnancy. *Br J Ind Med.* 1988;45:148–157.
8. Wilcox AJ, Horney LF. Accuracy of spontaneous abortion recall. *Am J Epidemiol.* 1984;120:727–733.
9. Bonney G. Logistic regression for dependent binary observations. *Biometrics.* 1987;43:951–973.
10. Bruzzi P, Green SB, Byar DP, Brinton LA, Schairer C. Estimating population attributable risk for multiple risk factors using case control data. *Am J Epidemiol.* 1985;122:904–914.
11. Leridon H. Facts and artifacts in the study of intra-uterine mortality: a reconsideration from pregnancy histories. *Popul Stud.* 1976;30:319–335.
12. Vaughan TL, Daling JR, Starzyk PM. Fetal death and maternal occupation. An analysis of birth records in the state of Washington. *J Occup Med.* 1984;26:676–678.
13. Roman E. Fetal loss rates and their relation to pregnancy order. *J Epidemiol Community Health.* 1984;38:29–35.
14. MacKenzie SG, Lippman A. An investigation of report bias in a case-control study of pregnancy outcome. *Am J Epidemiol.* 1989;129:65–75.
15. Wilcox AJ, Weinberg RC, and Baird DD. Risk factors for early pregnancy loss. *Epidemiol.* 1990;1:382–385.

Cigarette, Alcohol, and Coffee Consumption and Prematurity

Alison D. McDonald, MD, Ben G. Armstrong, PhD, and Margaret Sloan, BA

ABSTRACT

We analyzed data from a survey of occupational and other factors in pregnancy to assess the effects of cigarette, alcohol, and coffee consumption on pregnancy outcome. The risk of low birth weight for gestational age was found to increase substantially with smoking. Occasional consumers of alcohol had a slightly reduced risk relative to total abstainers. In more frequent drinkers, there was a small increase in risk. Risk increased slightly with coffee consumption. (*Am J Public Health.* 1992;82:87–90)

Introduction

The reduction in birth weight that accompanies maternal smoking in pregnancy, first reported in the 1950s,[1,2] has been well established.[3,4] Studies of alcohol[5–12] and caffeine[13–16] consumption have given conflicting findings. We used data from a survey in Montreal of occupational and other factors in pregnancy to examine the effects of smoking and consumption of alcohol and caffeine on pregnancy outcome. The survey is described briefly in the first of this series of papers and in detail in earlier reports.[17–20]

Methods

The present analysis was confined to current (just-completed) single pregnancies. Of the 51 512 pregnancies, 10 364 were excluded because of factors known to affect length of gestation and birth weight: multiple pregnancy (475), induction of labor (6113), hypertension (1187), diabetes (154), and prenatal bleeding or illness (2435). For another 362 women smoking, alcohol, or coffee consumption were unknown, and for 341 outcome information was incomplete, leaving 40 445 pregnancies for analysis.

To separate factors that affect fetal growth from those that diminish length of gestation, we defined prematurity in three ways: (1) low birth weight (≤2500 g), (2) preterm birth (<37 weeks), and (3) low birth weight for gestational age (LBWGA; bottom 5%).

The overall rates among the pregnancies surveyed were as follows: low birth

The authors are with the School of Occupational Health, McGill University, Montreal, Quebec.

Requests for reprints should be sent to Ben G. Armstrong, PhD, School of Occupational Health, McGill University, 1130 Pine Avenue West, Montreal, Quebec, Canada H3A 1A3.

This paper was submitted to the journal July 19, 1990, and accepted with revisions May 6, 1991.

TABLE 1—Risk of Prematurity by Maternal Characteristics

		Low Birth Weight		LBWGA		Preterm Birth	
	n	%	OR	%	OR	%	OR
Age							
< 20 years	1 627	8.2	1.00	8.1	1.00	10.3	1.00
20–29	27 557	5.8	1.10	5.4	1.13	6.7	0.74
30–34	8 661	4.8	1.18	4.3	1.28	6.7	0.79
35+	2 600	5.4	1.44	4.1	1.36	7.9	0.92
Pregnancy order							
1	19 096	6.0	1.00	6.2	1.00	6.4	1.00
2	12 909	5.2	0.67	4.3	0.60	7.1	0.93
3	5 485	5.6	0.66	4.2	0.60	8.0	0.95
4+	2 955	5.4	0.52	4.0	0.63	8.2	0.81
Previous spontaneous abortion							
None	34 425	5.4	1.00	5.2	1.00	6.6	1.00
One or more							
At 2nd pregnancy	1 771	8.4	2.13	6.5	1.73	8.7	1.51
At 3rd pregnancy	2 273	6.6	1.44	4.7	1.18	8.8	1.32
At pregnancy 4+	1 976	6.1	1.50	4.1	0.96	8.9	1.36
Previous low-birth-weight infant							
None	38 955	5.3	1.00	5.1	1.00	6.4	1.00
One or more							
At 2nd pregnancy	618	14.4	3.76	5.3	1.28	22.8	4.59
At 3rd pregnancy	463	14.7	3.16	6.3	1.33	21.8	3.81
At pregnancy 4+	409	13.2	3.06	5.9	1.31	19.3	3.28
Prepregnancy weight							
< 52 kg	15 339	8.1	1.00	7.7	1.00	8.0	1.00
52–58 kg	13 496	4.7	0.62	4.1	0.56	6.6	0.86
> 58 kg	11 610	3.6	0.47	3.1	0.42	6.0	0.76
Ethnic Group (color/language)							
White/English	7 221	3.3	1.00	3.5	1.00	5.3	1.00
White/French	25 959	6.1	1.44	5.6	1.21	7.0	1.21
Other	7 235	6.5	1.93	5.2	1.47	8.4	1.56
Education completed							
< Secondary	9 551	8.2	1.00	7.6	1.00	8.8	1.00
< College	17 387	5.7	0.84	5.1	0.80	7.2	0.92
College	13 507	3.8	0.69	3.5	0.68	5.3	0.72
Employment at start of pregnancy							
Unemployed	16 524	6.0	1.00	5.2	1.00	7.6	1.00
Employed	23 921	5.4	0.99	5.1	1.02	6.5	0.99
All	40 445	5.7		5.2		7.0	

Note. LBWGA = low birthweight for gestational age. OR = odds ratio, estimated by logistic regression with all factors, including cigarette, alcohol, and coffee consumption, in the model.

weight 5.7%, preterm birth 7.0%, LBWGA 5.2% (Table 1). To investigate the dependence of these three measures on the factors of interest, we used logistic regression, adjusting for the confounding factors shown in Table 1.

Results

Smoking

The risk of all three measures of prematurity, especially LBWGA, was higher among smokers than nonsmokers, and increased with number of cigarettes smoked ($P < .001$) (Table 2). For every 10 cigarettes smoked per day, the risk of LBWGA increased by a factor of 1.51 (95% CI, 1.44–1.57). Smoking accounted for 39% of cases of LBWGA, 35% of cases of low birth weight, and 11% of preterm births.

We investigated risk according to changes in smoking habits only for low birth weight. Women who smoked before but not during the first trimester had no excess risk (Table 3). Risk was also consistently reduced in women who cut down their consumption. Cessation before the second trimester brought the risk close to that for nonsmokers (Table 4). It is less clear that continued but reduced consumption was associated with a reduction in risk.

Mean placental weight varied little with smoking. For nonsmokers it was 626.6 g; for smokers of 1 to 9 cigarettes daily, 617.8 g; for 10 to 19 cigarettes, 618.7 g; and for 20 or more, 614.8 g. Accounting for birth weight, the placentas of smokers were a little heavier than those of nonsmokers; for example, the adjusted mean weight for heavy smokers was 30 g more than for nonsmokers.

Alcohol

For all three outcomes, light consumers of alcohol had a small but statistically significant reduction in risk relative to total abstainers (Table 5). Risk then appears to increase with alcohol intake, especially for low birth weight ($P = .002$ for trend). We examined risk of low birth weight by type of alcohol. After controlling for number of drinks, we found that risk was higher by a factor of 1.24 (95% CI, 1.00–1.53) for beer drinkers, and by a factor of 1.38 (95% CI, 0.96–1.99) for drinkers of spirits, than for wine drinkers.

Coffee

Trends of increasing risk with coffee consumption were consistent for low birth weight ($P = .02$) and LBWGA ($P = .01$), reaching a peak at a factor of about 1.4 in women consuming 10 or more cups per day (Table 6). Risk of LBWGA increased by a factor of 1.04 (95% CI, 1.01–1.06) per cup of coffee per day. If these associations were causal, coffee would account for 6% of cases of LBWGA, and 4% of cases of low birth weight.

Interactions

Interactions between smoking, alcohol, coffee, and other variables were seldom substantial but often statistically significant ($P < .05$). For LBWGA, the risk of smoking was apparently greater for French-speaking women than for others by a factor of 1.55 (95% CI, 1.17–2.05); that of alcohol consumption was 1.35 times greater for mothers over 30 (95% CI, 1.08–1.69); that of coffee consumption was 1.27 times greater for alcohol consumers than for abstainers (95% CI, 1.02–1.60).

Discussion

A reduction in birth weight was found that was consistent with previous reports that infants of smokers weighed on average 150 to 250 g less than those of nonsmokers.[3] There was a quantitative relationship with the number of cigarettes smoked. Our exploration of interactions indicated that this relationship was not importantly influenced by any other factor. A slightly shorter gestation period, a mat-

ter of a day or two, was found in smokers in our study, as in others. The somewhat greater effect of smoking on birth weight corrected for gestation length than on uncorrected birth weight indicates that smoking directly retards fetal growth. The higher ratio of placental to infant weight that we found in smokers has also been observed by others.[21–23] Our finding that women who stopped smoking during the first trimester had little or no greater risk of having a low-birth-weight infant than did nonsmokers, together with other similar findings,[9,24,25] is important.

In contrast, the evidence for alcohol as a risk factor was unconvincing. Women who took fewer than seven drinks a week had fewer low-birth-weight babies than did those who drank more than this or those who drank no alcohol. This result is consistent with the finding of higher average birth weights in babies born to women who drank alcohol than in those who did not.[11] Two studies have been reported, one in Britain[7] and another in the United States,[8] in which retarded fetal growth was significantly more frequent when 10 or more drinks a week were taken. In France, an increase in small-for-dates babies was found among women who drank three or more drinks a day, but only in women who were mainly beer drinkers.[5] In our study also, such risk as there was predominantly affected beer and spirits drinkers. Our study and two others[9,26] suggest that there is no effect on birth weight when fewer than two or three drinks a day are taken, and that there is at most a small effect at higher levels of consumption. This conclusion remains compatible with the positive studies.[6,7]

The small increase in risk, with evidence of dose-response, suggests that heavy coffee consumption may have a small effect on birth weight. This finding is compatible with reported increases in risks by a factor of 1.24 for heavy coffee consumption[8] and by a factor of 1.58 for frequent coffee consumption.[16] In a study of total caffeine consumption, however, a fourfold increase in risk of low birth weight was found with the consumption of caffeine equivalent to more than three cups of coffee a day.[14] In Quebec women in the Canada Nutrition Survey, total caffeine consumption and coffee consumption were highly correlated (r = .98) (Health and Welfare Canada, personal communication). The difference between these two measures is thus unlikely to explain the higher risk found by the latter[14] study. □

TABLE 2—Risk of Prematurity by Cigarette Consumption

	No. Pregnancies	% Premature	OR	95% CI	Change[a]
Low birth weight					
Nonsmoker	26 089	3.6	1.00		73.4 g
< 10 per day	3 147	6.1	1.64	1.39–1.93	−37.1 g
10–19 per day	5 042	9.2	2.39	2.12–2.71	−138.3 g
20+ per day	6 167	11.2	2.85	2.53–3.21	−178.7 g
LBWGA					
Nonsmoker	26 089	3.0	1.00		59.8 g
< 10 per day	3 147	6.3	1.97	1.68–2.33	−29.5 g
10–19 per day	5 042	8.6	2.58	2.27–2.93	−107.2 g
20+ per day	6 167	10.8	3.19	2.82–3.60	−150.4 g
Preterm birth					
Nonsmoker	26 089	6.0	1.00		−0.08 wk
< 10 per day	3 147	7.4	1.22	1.05–1.41	−0.08 wk
10–19 per day	5 042	9.1	1.43	1.27–1.60	−0.20 wk
20+ per day	6 167	8.9	1.33	1.18–1.49	−0.16 wk

Note. OR = odds ratio, estimated by logistic regression with all factors, including age, pregnancy order, previous spontaneous abortion, previous low-birth-weight infant, prepregnancy weight, ethnic group, education employment at start of pregnancy, and alcohol and coffee consumption, in the model. CI = confidence interval. LBWGA = low birth weight for gestational age.
[a]Mean, adjusted for other risk factors by analysis of variance.

TABLE 3—Risk of Low Birth Weight by Changes in Smoking Habit at the Start of Pregnancy

Cigarettes per Day before Pregnancy	Cigarettes per Day in First Trimester			
	0	1–9	10–19	20+
0	1	—	—	—
1–9	0.64 (0.37–1.10)	1.15 (0.79–1.66)	—	—
10–19	0.98 (0.65–1.50)	1.81 (1.34–2.45)	2.20 (1.80–2.68)	—
20+	0.98 (0.70–1.39)	2.18 (1.60–2.96)	2.67 (2.40–3.38)	3.05 (2.66–3.50)

Note. Figures given in the table are odds ratios adjusted for confounders, relative to women smoking in neither period. 95% confidence intervals are given in parentheses.

TABLE 4—Risk of Low Birth Weight by Changes in Smoking Habits between the First and Second Trimester

Cigarettes per Day in First Trimester	Cigarettes per Day in Second Trimester			
	0	1–9	10–19	20+
0	1	—	—	—
1–9	0.95 (0.53–1.71)	1.05 (0.58–1.89)	—	—
10–19	1.22 (0.68–2.21)	2.98 (1.87–4.75)	2.70 (2.32–3.14)	—
20+	1.19 (0.68–2.10)	2.34 (1.21–4.53)	3.43 (2.60–4.53)	3.26 (2.83–3.75)

Note. Figures given in the table are odds ratios adjusted for confounders, relative to women smoking in neither period. 95% confidence intervals are given in parentheses.

References

1. Simpson WJ. A preliminary report of cigarette smoking and the incidence of prematurity. *Am J Obstet Gynecol.* 1957;73:808–815.
2. Lowe CR. Effect of mothers' smoking habits on birth weight of their children. *Br Med J.* 1959;2:673–676.
3. United States Surgeon General. *The Health Consequences of Smoking: A Report of the Surgeon-General.* Washington, DC: US Govt Printing Office; 1973. Dept of Health, Education, and Welfare publication HSM 73-8704.
4. American Academy of Pediatrics Committee on Environmental Hazards. *Effects of cigarette smoking on the fetus and child. Pediatr.* 1976;57:411–413.
5. Kaminski M, Rumean-Rouquette C, Schwartz D. Consommation d'alcool chez les femmes enceintes et issue de grossesse. *Rev Epidemiol Santé Publique.* 1976; 16:27–40.

TABLE 5—Risk of Prematurity by Alcohol Consumption

	No. Pregnancies	% Premature	OR	95% CI	Change[a]
Low Birth Weight					
None	27 804	6.0	1.00		−10.2 g
1–2 drinks per week	7 330	4.4	0.79	0.70–0.90	21.3 g
3–6 drinks per week	3 970	5.0	0.87	0.74–1.01	24.8 g
7–20 drinks per week	1 203	6.3	1.03	0.80–1.31	24.9 g
21+ drinks per week	138	11.6	1.59	0.92–2.73	0.5 g
LBWGA					
None	27 804	5.5	1.00		−5.8 g
1–2 drinks per week	7 330	4.1	0.80	0.70–0.91	11.6 g
3–6 drinks per week	3 970	4.5	0.85	0.72–1.00	12.9 g
7–20 drinks per week	1 203	6.0	1.02	0.79–1.31	19.0 g
21+ drinks per week	138	8.7	1.19	0.65–2.19	20.2 g
Preterm Birth					
None	27 804	7.3	1.00		−0.04 wk
1–2 drinks per week	7 330	5.8	0.86	0.77–0.96	0.07 wk
3–6 drinks per week	3 970	6.2	0.90	0.79–1.04	0.10 wk
7–20 drinks per week	1 203	7.0	0.99	0.78–1.24	0.09 wk
21+ drinks per week	138	9.4	1.21	0.68–2.17	−0.01 wk

Note. OR = odds ratio, estimated by logistic regression with all factors, including age, pregnancy order, previous spontaneous abortion, previous low-birth-weight infant, prepregnancy weight, ethnic group, education, employment at start of pregnancy, and cigarette and coffee consumption in the model. CI = confidence interval. LBWGA = low birth weight for gestational age.
[a]Mean, adjusted for other risk factors by analysis of variance.

TABLE 6—Risk of Prematurity by Coffee Consumption

	No. Pregnancies	% Premature	OR	95% CI	Change[a]
Low Birth Weight					
None	17 518	5.0	1.00		16.3 g
1–2 cups per day	7 471	5.5	1.05	0.95–1.16	−4.0 g
3–4 cups per day	3 743	7.5	1.08	0.93–1.25	−33.7 g
5–9 cups per day	1 380	9.5	1.13	0.92–1.39	−46.2 g
10+ cups per day	333	12.9	1.43	1.02–2.02	−81.9 g
LBWGA					
None	17 518	4.5	1.00		16.8 g
1–2 cups per day	17 471	4.9	1.05	0.94–1.16	−3.9 g
3–4 cups per day	3 743	7.2	1.15	0.99–1.34	−37.3 g
5–9 cups per day	1 380	10.0	1.34	1.10–1.65	−46.0 g
10+ cups per day	333	11.4	1.39	0.97–1.98	−73.0 g
Preterm Birth					
None	17 518	6.7	1.00		0.00 wk
1–2 cups per day	17 471	6.7	1.00	0.92–1.09	0.00 wk
3–4 cups per day	3 743	8.2	1.08	0.94–1.24	0.03 wk
5–9 cups per day	1 380	8.5	1.06	0.86–1.30	−0.01 wk
10+ cups per day	333	10.5	1.24	0.86–1.79	−0.06 wk

Note. OR = odds ratio, estimated by logistic regression with all factors, including age, pregnancy order, previous spontaneous abortion, previous low-birth-weight infant, prepregnancy weight, ethnic group, education, employment at start of pregnancy, and cigarette and alcohol consumption in the model. CI = confidence interval. LBWGA = low birth weight for gestational age.
[a]Mean, adjusted for other risk factors by analysis of variance.

6. Mills JL, Granbard BI, Harley EE, Rhoads GG, Berendes HW. Maternal alcohol consumption and birth weight. *JAMA.* 1984; 252:1895–1899.
7. Wright JT, Barrison IG, Lewis IG, et al. Alcohol consumption, pregnancy and low birth weight. *Lancet.* 1983;1:663–665.
8. Van den Berg BJ. Epidemiologic observations of prematurity: effects of tobacco, coffee and alcohol. In: Reed DM, Stanley FJ, eds. *Epidemiology of Prematurity.* Baltimore, Md: Urban and Schwarzenberg; 1977:157–176.
9. Kline J, Stein Z, Hutzler M. Cigarettes, alcohol and marijuana: varying associations with birth weight. *Int J Epidemiol.* 1987;16:44–52.
10. Grisso JA, Roman E, Inskip H, Beral V, Donovan J. Alcohol consumption and outcome of pregnancy. *J Epidemiol Community Health.* 1984;38:232–235.
11. Brooke OG, Anderson HR, Bland JM, Peacock JL, Stewart CM. Effect on birthweight of smoking, alcohol, caffeine, socioeconomic factors and psychosocial stress. *Br Med J.* 1989;298:795–801.
12. Walpole I, Zubrick S, Pontré J. Is there a fetal effect with low to moderate alcohol use before or during pregnancy? *J Epidemiol Community Health.* 1990;44:297–301.
13. Mau G, Netter P. Are coffee and alcohol consumption risk factors in pregnancy. *Geburtshilfe und Frauenheitkunde.* 1974;34:1018–1022.
14. Martin TR, Bracken MB. The association between low birth weight and caffeine consumption during pregnancy. *Am J Epidemiol.* 1987;126:813–821.
15. Linn S, Schoenbaum SC, Monson RR, Rosner B, Stubblefield PG, Ryan KJ. No association between coffee consumption and adverse outcome of pregnancy. *N Engl J Med.* 1982;306:141–145.
16. Hogue CJ. Coffee in pregnancy. *Lancet.* 1981;1:554.
17. McDonald AD, McDonald JC, Diodati-Nolin A, et al. Occupation and pregnancy outcome. *Br J Ind Med.* 1987;44:521–526.
18. McDonald AD. Work and pregnancy. *Br J Ind Med.* 1988;45:577–580.
19. McDonald AD, McDonald JC, Armstrong BG, Cherry NM, Nolin A, Robert D. Prematurity and work in pregnancy. *Br J Ind Med.* 1988;45:56–62.
20. Armstrong BG, Nolin AD, McDonald AD. Work in pregnancy and birth weight for gestational age. *Br J Ind Med.* 1989;46:196–199.
21. Mulcahy R, Murphy J, Martin F. Placental changes and maternal weight in smoking and non-smoking mothers. *Am J Obstet Gynecol.* 1970;106:703–704.
22. Wyngerd J, Christianson R, Levitt WV, Schoen EJ. Placental ratio in white and black women: relation to smoking and anaemia. *Am J Obstet Gynecol.* 1976;124:671–675.
23. Naeye RL. Effects of maternal cigarette smoking on the fetus and placenta. *Br J Obstet Gynaecol.* 1978;85:732–737.
24. MacArthur C, Knox EG. Smoking in pregnancy: effect of stopping at different stages. *Br J Obstet Gynaecol.* 1988;95:551–555.
25. Sexton M, Habel JR. A clinical trial of change in maternal smoking and its effect on birth weight. *Am J Public Health.* 1984;251:911–915.
26. Marbury MC, Linn S, Monson RR, Schoenbaum S, Stubblefield PG, Ryan KJ. The association of alcohol consumption with outcome of pregnancy. *Am J Public Health.* 1983;73:1165–1168.

Cigarette, Alcohol, and Coffee Consumption and Congenital Defects

Alison D. McDonald, MD, Ben G. Armstrong, PhD, and Margaret Sloan, BA

ABSTRACT

We analyzed data from a survey of occupational and other factors and pregnancy outcome to assess the effects of cigarette, alcohol, and coffee consumption. There was no evidence of an association between any congenital defect and smoking. Results for alcohol and coffee consumption were largely negative, but there was a weak association with musculoskeletal defects in babies born to women who drank one or more alcoholic drinks a day. (*Am J Public Health.* 1992;82:91–93)

Introduction

A report that smoking in pregnancy was associated with congenital heart defects[1] was not confirmed,[2-4] even in two large series of births.[5] Heavy alcohol consumption has been reported to cause retarded physical and mental development and certain facial and neurological abnormalities (fetal alcohol syndrome). Moderate alcohol consumption has also been reported to affect fetal growth and morphogenesis.[6] Some concern has been expressed that heavy coffee consumption may be teratogenic,[7] although a case-control study in Finland was negative.[8] We investigated the relationships between cigarette, alcohol, and coffee consumption and congenital defects, using data from a survey of occupational factors in pregnancy conducted in Montreal from 1982 to 1984 (see the first paper in this series). In this survey no association was found between congenital defects and chemical exposures estimated from job title,[9] but in a case-referent study an association with aromatic solvents was found.[10] The excess defects were mainly renal and urinary. The association of cigarette, alcohol, and coffee consumption with congenital defects has now been analyzed for all current and previous pregnancies of the 56 000 women interviewed—104 000 pregnancies in all.

Methods

The study design and methods are briefly described in the first of these reports, and have been described in detail elsewhere.[11] All current and previous pregnancies of more than 20 weeks' duration, whether or not the woman was employed during the pregnancy, were included in the present analysis. Therapeutic abortions for fetal defect were included irrespective of gestation length. Information on cigarette and alcohol consumption, and on potential confounding factors was available for 89 317 of these pregnancies. As questions on coffee consumption in previous pregnancies were introduced after the survey began, only 80 319 remained for study of this factor.

Congenital defects were ascertained (blind) in current pregnancies from pediatric records.[8] For previous pregnancies, the mother was asked whether any abnormality had been noted in the child. If a defect was suspected, further inquiries were made. The defects were classified into 90 groups. Minor abnormalities such as hydrocele, glandular hypospadias, undescended testes, skin tags, nevi, and hernias were omitted, as were known inherited defects, those that were poorly described (less than 1%), and those that became apparent only at a later stage of development, such as mental retardation, epilepsy, cerebral palsy, and deafness. Interauricular and interventricular septal defects were retained if accompanied by another heart abnormality. Cases with multiple defects were classified under the one that appeared to be most serious. Eight types of congenital defect (shown in Tables 1 and 2) were entered into the present analyses.

Possible confounding by maternal age, educational level, and ethnic (color and language) group was controlled by including these factors together with cigarette, alcohol, and coffee consumption in logistic regression analyses, which were performed separately for current and previous pregnancies. Regression coefficients for cigarette, alcohol, and coffee consumption from the two series were combined as a weighted mean to obtain a single odds ratio, as they did not significantly differ for any factor at a 5% probability level.

Results

There was no consistent evidence of an association between cigarette consumption and any defect group, after

The authors are with the School of Occupational Health, McGill University, Montreal, Quebec.

Requests for reprints should be sent to Ben G. Armstrong, PhD, School of Occupational Health, McGill University, 1130 Pine Avenue West, Montreal, Quebec, Canada H3A 1A3.

This paper was submitted to the journal July 19, 1990, and accepted with revisions May 6, 1991.

TABLE 1—Congenital Defects by Type According to Cigarette Consumption in the First Trimester

Defect	Nonsmoker[a] n	Cigarettes per Day: 1 to 9 n	1 to 9 OR (95% CI)	10 to 19 n	10 to 19 OR (95% CI)	20+ n	20+ OR (95% CI)	Total n	Total Rate[b]
Chromosomal	123	12	1.02 (0.7–1.5)	13	0.94 (0.7–1.3)	21	0.84 (0.6–1.2)	169	1.9
Neural tube	106	21	1.84 (1.1–3.0)	23	1.06 (0.7–1.7)	40	1.35 (0.9–2.0)	190	2.1
Cleft of lip or palate	59	8	1.17 (0.6–2.5)	11	1.04 (0.5–2.1)	18	1.09 (0.6–1.9)	96	1.1
Cardiovascular	200	20	1.03 (0.6–1.6)	41	1.17 (0.8–1.7)	57	1.17 (0.9–1.6)	318	3.6
Digestive or respiratory	69	6	0.75 (0.3–1.7)	16	1.13 (0.7–2.0)	18	0.94 (0.5–1.7)	109	1.2
Renal or urinary	124	20	1.55 (1.0–2.5)	32	1.42 (0.9–2.1)	33	1.17 (0.8–1.8)	209	2.3
Clubfoot	396	47	1.06 (0.8–1.4)	76	1.00 (0.8–1.3)	95	0.98 (0.8–1.3)	614	6.9
Musculoskeletal	149	17	1.07 (0.6–1.8)	33	1.29 (0.9–1.9)	24	0.73 (0.5–1.2)	223	2.5
Total	1 226	151	1.14 (1.0–1.4)	245	1.08 (0.9–1.2)	306	1.02 (0.9–1.2)	1 928	21.6
No. pregnancies	59 456	6 308		10 304		13 249		89 317	

Note. n = number of defects. OR = odds ratio, accounting for confounding by education, ethnicity, maternal age, and alcohol and coffee consumption. CI = confidence interval.
[a]OR = 1.00.
[b]Rate per 1000 pregnancies.

TABLE 2—Congenital Defects by Type According to Alcohol Consumption in the First Trimester

Defect	None[a] n	Drinks per Week: 1 to 2 n	1 to 2 OR (95% CI)	3 to 6 n	3 to 6 OR (95% CI)	7+ n	7+ OR (95% CI)	Total n	Total Rate[b]
Chromosomal	110	40	1.19 (0.9–1.5)	16	1.11 (0.8–1.5)	3	0.91 (0.5–1.6)	169	1.9
Neural tube	139	39	1.19 (0.8–1.7)	8	0.59 (0.3–1.2)	8	0.63 (0.2–1.7)	190	2.1
Cleft of lip or palate	62	21	1.42 (0.9–2.4)	11	1.68 (0.9–3.3)	2	1.06 (0.3–4.2)	96	1.1
Cardiovascular	225	55	0.91 (0.7–1.2)	24	0.96 (0.6–1.5)	14	1.24 (0.7–2.2)	318	3.6
Digestive or respiratory	70	23	1.27 (0.8–2.1)	12	1.56 (0.8–2.9)	4	2.00 (0.7–5.6)	109	1.2
Renal or urinary	152	36	0.86 (0.6–1.2)	19	1.02 (0.6–1.7)	2	0.59 (0.1–2.4)	209	2.3
Clubfoot	412	130	1.19 (1.0–1.5)	53	1.03 (0.8–1.4)	19	0.99 (0.6–1.6)	614	6.9
Musculoskeletal	143	44	1.13 (0.8–1.6)	23	1.33 (0.8–2.1)	13	1.85 (1.0–3.3)	223	2.5
Total	1 313	388	1.11 (1.0–1.3)	166	1.06 (0.9–1.3)	61	0.95 (0.7–1.2)	1 928	21.6
No. pregnancies	62 539	16 441		7357		3000		89 317	

Note. n = number of defects. OR = odds ratio, accounting for confounding by education, ethnicity, maternal age, and cigarette and coffee consumption. CI = confidence interval.
[a]OR = 1.00.
[b]Rate per 1000 pregnancies.

accounting for possible confounding (Table 1). Odds ratios (relative to nonsmokers) for mothers who smoked one to nine cigarettes per day were 1.84 for neural tube, 1.53 for renal-urinary, and 1.14 for all defects. These values were on the borderline of statistical significance (at about $P = .05$); however, odds ratios were lower in heavier smokers.

There was some evidence of an increased risk of musculoskeletal defects in women who took alcohol, with risk increasing with the amount drunk (Table 2). The odds ratio of 1.82 in women who took seven or more drinks per week was significant at $P = .05$. Among the 13 defects in this group were 6 cases of congenital dislocation of the hip, whereas 2.3 were expected based on the proportion of musculoskeletal defects that were of this type overall. The specific risk for congenital dislocation of the hip (6/2.3 = 2.61) was thus somewhat higher than for musculoskeletal defects overall, but because of small numbers it is not statistically significant ($P > .05$). The other 7 musculoskeletal defects were craniosynostosis, scaphocephaly, cranial abnormality, hemivertebrae, limb reduction deformity (2), and fibular abnormality. Facial deformities were not observed.

Of the odds ratios for defects in babies born to coffee drinkers (Table 3), only that for heart defects among the children of women who drank three or more cups a day (1.52) was significantly elevated ($P = .02$). No specific type of heart defect was overrepresented in these 58 cases compared with defects in babies born to women who did not drink coffee.

TABLE 3—Congenital Defects by Type According to Coffee Consumption in the First Trimester

Defect	Cups per Day: None[a] n	1 to 2 n	1 to 2 OR (95% CI)	3+ n	3+ OR (95% CI)	Total n	Total Rate[b]
Chromosomal	61	69	1.01 (0.8–1.2)	20	0.97 (0.7–1.3)	150	1.9
Neural tube	63	71	1.12 (0.8–1.6)	28	1.09 (0.7–1.8)	162	2.0
Cleft of lip or palate	29	43	1.42 (0.9–2.3)	15	1.39 (0.7–2.7)	87	1.1
Cardiovascular	101	112	1.08 (0.8–1.4)	58	1.52 (1.1–2.2)	271	3.4
Digestive or respiratory	35	44	1.17 (0.7–1.8)	18	1.34 (0.7–2.4)	97	1.2
Renal or urinary	83	83	0.96 (0.7–1.3)	25	0.81 (0.5–1.3)	191	2.4
Clubfoot	228	248	1.01 (0.8–1.2)	76	0.91 (0.7–1.2)	552	6.9
Musculoskeletal	82	78	0.89 (0.6–1.2)	36	1.30 (0.9–2.0)	197	2.4
Total	682	748	1.04 (0.9–1.2)	276	1.09 (0.9–1.3)	1 706	21.2
No. pregnancies	34 282	34 345		11 691		80 319	

Note. n = number of defects. OR = odds ratio, accounting for confounding by education, ethnicity, maternal age, and cigarette and alcohol consumption. CI = confidence interval.
[a]OR = 1.00.
[b]Rate per 1000 pregnancies.

Discussion

Given that over 70 odds ratios were tested, the three that were significantly ($P < .05$) raised could have occurred by chance. The one association with smoking, namely an increase of neural tube defects in babies born to women who smoked 1 to 9 cigarettes per day, would have been more convincing if an excess had been seen in the heavier-smoking groups. Even quite heavy cigarette smoking was common in this survey—in 30% of the pregnancies the mother smoked during the first trimester of pregnancy, and in 15% the mother smoked 20 or more cigarettes per day. Thus our study had the power to detect important risks due to smoking, if such risks were present. The risks we can exclude are specified by the upper confidence limits of the odds ratios, which are quite low for smoking.

Alcohol consumption, on the other hand, although perhaps underreported, was generally low. Only 3% of the women surveyed reported taking one or more drinks per day (seven per week), and 0.3% reported taking three or more drinks per day. The power of our study to detect increased risk at these levels was thus very limited, as is evident from the wide confidence limits for the odds ratios shown in Table 2. No case of suspected fetal alcohol syndrome was noted, although we were not specifically looking for the syndrome. The association of alcohol with musculoskeletal defects has not been reported elsewhere, and was of low statistical significance.

Coffee consumption was associated only with heart defects, and the evidence was not strong. In Finland, where coffee consumption is high (29% of those surveyed reported taking three or more cups of coffee a day, compared with 12% in Montreal), no association was found between coffee consumption and clefts of the lip and palate or defects of the central nervous, cardiovascular, or musculoskeletal systems. □

References

1. Fredrick J, Alberman E, Goldstein H. Possible teratogenic effect of cigarette smoking. *Nature.* 1971;231:529–530.
2. Comstock GW, Shah FK, Meyer MB, Abbey H. Low birth weight and neonatal mortality rate related to maternal smoking and socioeconomic status. *Am J Obstet Gynecol.* 1971;111:53–59.
3. Kullander S, Kallen, B. A prospective study of smoking and pregnancy. *Acta Obstet Gynecol Scand.* 1971;50:83–94.
4. Andrews J, McGarry JM. A community study of smoking in pregnancy. *J Obstet Gynecol BC.* 1972;79:1057–1073.
5. Shiono PH, Klebanoff MA, Berendes HW. Congenital malformations and maternal smoking during pregnancy. *Teratology.* 1986;34:65–71.
6. Hanson J, Streissguth A, Smith D. The effects of moderate alcohol consumption during pregnancy on fetal growth and morphogenesis. *J Pediatr.* 1978;92:457–460.
7. Soyka L. Caffeine ingestion during pregnancy: in utero exposure and possible effects. *Semin Perinatol.* 1981;5:305–309.
8. Kurppa K, Holmberg P, Kuosmer E, Saxen L. Coffee consumption during pregnancy and selected congenital malformations: a nationwide case-control study. *Am J Public Health.* 1983;73:1397–1399.
9. McDonald AD, McDonald JC, Armstrong B, et al. Congenital defects and work in pregnancy. *Br J Ind Med.* 1988;45:581–588.
10. McDonald JC, Lavoie J, Côté R, McDonald AD. Chemical exposures at work in early pregnancy and congenital defect: a case-referent study. *Br J Ind Med.* 1987;44:527–533.
11. McDonald AD, McDonald JC, Armstrong B, et al. Occupation and pregnancy outcome. *Br J Ind Med.* 1987;44:521–526.

The Effect of State Cigarette Tax Increases on Cigarette Sales, 1955 to 1988

Dan E. Peterson, MD, MPH, Scott L. Zeger, PhD, Patrick L. Remington, MD, MPH, and Henry A. Anderson, MD

ABSTRACT

We evaluated the effect of state cigarette tax increases on cigarette sales in the 50 states for the years 1955 to 1988. State cigarette tax increases were associated with an average decline in cigarette consumption of three cigarette packs per capita (about 2.4%). Larger tax increases were associated with larger declines in consumption. Raising state cigarette taxes appears to be an effective public health intervention that can reduce cigarette consumption and its associated health consequences. (*Am J Public Health.* 1992; 82:94–96)

Introduction

Numerous studies have observed that cigarette consumption decreases when the price of cigarettes increases.[1–6] Although cigarette tax increases have generally been enacted to raise revenues, such tax increases have raised the price of cigarettes and are believed to have led to reductions in cigarette consumption. There was a substantial drop in national cigarette sales in 1983 following the increase of the federal cigarette excise tax from 8 to 16 cents on January 1, 1983. Several observers have commented on this as a cause–effect relationship.[7–11] It is likely that state cigarette tax increases, of which there have been 248 during the years 1955 through 1988, have had similar effects. In this context, we undertook the present study to assess changes in cigarette consumption following state cigarette tax increases and to quantify the amount of the change in cigarette consumption by the size of tax increase.

Methods

Annual data on cigarette packs per capita (number of packs sold in the state divided by the state's population), average retail price in the state, and state cigarette taxes were obtained from the Tobacco Institute.[12] Data were available from 1955 or when a state first imposed a cigarette tax, whichever was later. We converted all tax and price data to constant (1983) dollars.[13] Data on the enactment of state clean indoor air acts were obtained from the surgeon general's 1989 report.[7]

We defined the change in cigarette consumption for each state and each year as the difference between the sales in that state in the preceding year and the sales in the following year. Calculating the differences in sales for each of the states for each year of the study period gave us 1536 observations of changes in cigarette consumption. We adjusted for national trends in cigarette consumption by subtracting the national change in consumption for that year from each state value. This has the effect of controlling for national events, such as the 1983 federal tax hike, as well as national trends in sales due to such factors as national demographic trends and antismoking publicity.

To explore the effect of state tax increases on cigarette consumption, we compared changes in cigarette consumption for states in years with tax increases against states in years with no tax increase. We then categorized observations by the size of the tax increase and compared changes in cigarette consumption. We also categorized observations by the presence or absence of a new clean indoor air act and compared changes in cigarette consumption.

Means and standard errors of the categorized changes in cigarette consumption were computed. The standard errors were calculated using pooled estimates of the variance across subgroups, accounting for correlation among repeated events within a given state using the robust variance method of Zeger and Liang.[14]

After analyzing the data for the entire period (1955 through 1988), we looked at the effects in each of three time intervals: 1955 through 1964 (until the surgeon general's first report on smoking), 1965 through 1978 (when per-capita cigarette

Patrick L. Remington and Henry A. Anderson are with the Section of Environmental and Chronic Disease Epidemiology, Wisconsin Department of Health and Social Services, Madison, WI. Scott L. Zeger is with the Department of Biostatistics, Johns Hopkins School of Hygiene and Public Health, Baltimore. Dan E. Peterson is with the Division of Field Epidemiology, Epidemiology Program Office, Centers for Disease Control, Atlanta, Georgia. At the time of this work, Dr. Peterson was assigned to the Wisconsin Department of Health and Social Service.

Requests for reprints should be sent to Dan E. Peterson, MD, MPH, Division of Field Epidemiology, MS C08, Centers for Disease Control, Atlanta, GA 30333.

This paper was submitted to the journal April 23, 1991, and accepted with revisions September 20, 1991.

Editor's Note. See related editorial by Wasserman on page 19.

TABLE 1—Change in Cigarette Consumption by State Tax and State Clean Indoor Air Act, United States, 1955–1988

	Change in Consumption (cigarette packs per capita)			
	1955–1988	1955–1964	1965–1978	1979–1988
Tax increase				
Yes	−3.0[a]	−3.3[a]	−3.8[a]	−1.5[a]
No	0.6[a]	0.7	0.7	0.3
Size of tax increase				
<4.0 cents	0.8	−0.9	6.9[a]	−0.6
>4.0–6.0 cents	−1.5	−3.8	−1.5	−1.2
>6.0–8.0 cents	−4.6[a]	−4.7[a]	−5.0[a]	−3.8[a]
>8.0 cents	−7.3[a]	−8.7[a]	−7.4[a]	−3.1
Clean indoor air act				
Yes	−0.1	...	−1.5	0.1
No	0.0	...	0.3	0.0

Note. Average capita consumption nationally for the period was 124.1 packs per capita (1955–1964, 119.5; 1965–1978, 126.6; 1979–1988, 125.0).
[a]95% confidence interval does not include zero

consumption peaked), and 1979 through 1988 (a period of declining consumption).

Results

Cigarette consumption declined by an average of 3.0 packs per capita in the 248 observations with a tax increase. In contrast, cigarette consumption rose by 0.6 packs per capita in the 1288 observations with no tax increase (Table 1). Stratification of the change in cigarette consumption by the size of tax increase yielded a clear dose–response relationship: the larger the increase in taxes, the larger the decrease in consumption (Table 1).

Regressing the size of tax increase on the change in cigarette consumption, we found that for each 1-cent increase in cigarette taxes, sales fell by 0.75 packs per capita (about 0.6%). Performing a separate regression on the data for each of the three time periods, we found that while in each period the effect of tax on sales was consistent and dose related, it appears that a given size of tax increase had less effect in recent years (sales fell by 0.85, 0.78, and 0.48 pack per penny tax increase for the time periods 1955 through 1964, 1965 through 1978, and 1979 through 1988, respectively).

No consistent relationship was found between the enactment of state clean indoor air acts and changes in cigarette consumption.

Discussion

Our results are consistent with the econometric literature on cigarettes, in which estimated price elasticity for cigarettes (the percentage change in sales as a result of a 1% change in price) has ranged from −0.14 to −1.44 depending on the methods used and the populations studied. Reviewing this literature, the surgeon general's 1989 report used −0.47 as an estimate of the price elasticity for the general population.[7] In our study, using a rough conversion from our tax-response format to an elasticity format, we found an elasticity of −0.49 (in this case the elasticity representing a sales decline of 0.49% for every 1% of cigarette price that the tax increase accounted for).

Despite the strong and consistent relationship observed, it could be argued that changing public opinion regarding smoking precedes and causes both a tax increase and a decrease in cigarette sales. This hypothesis seems unlikely, as we found the same relationship between tax increases and declines in cigarette sales even in the first period of the analysis, 1955 through 1964, a period before the surgeon general's first report and a time when little antismoking sentiment existed. Additionally, the enactment of a clean indoor air act, a similar legislative event that would seemingly also reflect such an underlying public opinion shift, was not found to be related to changes in cigarette consumption.

Other studies[4,5] have found some evidence for a relationship between enactment of a state clean indoor air act and reductions in cigarette consumption. We conclude, from our analysis, that if cigarette consumption declines occurred at the time of enactment of state clean indoor air acts, they were small compared with those seen with tax increases. Alternatively, state clean indoor air acts may have a delayed and more gradual effect on cigarette consumption, and may therefore have escaped our detection.

While our approach is conceptually easy to understand, it is limited in that we were not able to control for potential biases, such as state-to-state differences in sales or price trends, or to examine two effects simultaneously, such as the interaction of taxes and clean indoor air acts. To evaluate whether these limitations were an important factor in our results, we did a second analysis, using a longitudinal model,[15,16] in which we were able to control for state-specific trends over time and to examine the role of multiple variables simultaneously. The results of this second analysis were essentially the same as the first analysis. Restricting the analysis to 21 states reported to have little or no problem with bootlegging[17] (cigarettes purchased in one state and consumed in another) also did not change the results.

We have used an epidemiologic approach to evaluate the effect of a state-based public health intervention—the enactment of cigarette excise tax increases. Our results provide further evidence that state tax increases are effective in reducing smoking and that the larger the tax increase, the larger the drop in cigarette smoking. □

Acknowledgments

This paper was presented at the Epidemic Intelligence Service 1990 Conference, Atlanta, April 1990; the International Epidemiologic Association XIIth International Conference, Los Angeles, August 1990; and the Fifth National Conference on Chronic Disease Control, Detroit, October 1990.

We express our appreciation to Kenneth Warner, PhD, Daniel Sosin, MD, James Marks, MD, and Joe Schirmer, MS, for their reviews and contributions to this paper.

References

1. Lyon HL, Simon JL. Price elasticity of the demand for cigarettes in the United States. *Am J Agricultural Economics*. 1968;50:888–895.
2. Lewit EM, Coate D, Grossman M. The effects of government regulation on teenage smoking. *Law and Economics*. 1981;24:545–569.
3. Lewit EM, Coate D. The potential for using excise taxes to reduce smoking. *Health Economics*. 1982;1:121–145.
4. Chaloupka F, Saffer H. *The Demand for Cigarettes and Restrictions on Smoking in the Workplace*. New York, NY: National Bureau of Economic Research; 1988. Working paper.
5. Chaloupka FJ IV. *An Economic Analysis of Addictive Behavior: The Case of Cigarette Smoking*. New York, NY: City University of New York; 1988. Doctoral dissertation.

6. Baltagi BH, Goel RK. Quasi-experimental price elasticities of cigarette demand and the bootlegging effect. *Am J Agricultural Economics.* 1987;69:750–754.
7. US Department of Health and Human Services. *Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General.* Washington, DC: US Department of Health and Human Services; Public Health Service; Centers for Disease Control; Center for Chronic Disease Prevention and Health Promotion; Office on Smoking and Health; 1989; DHHS publication no. (CDC) 89-8411.
8. Warner KE. Cigarette taxation: doing good by doing well. *J Public Health Policy.* 1984;3:312–319.
9. Warner KE. Smoking and health implications of a change in the federal cigarette excise tax. *JAMA.* 1986;255:1028–1032.
10. Harris JE. The 1983 increase in the federal cigarette excise tax. In: Summers LH, ed. *Tax Policy and the Economy.* Cambridge, Mass: Massachusetts Institute of Technology Press; 1987;1:87–111.
11. Lewit EM. US tobacco taxes: behavioural effects and policy implications. *Br J Addict.* 1989;84:1217–1235.
12. Tobacco Institute. *The Tax Burden on Tobacco, Historical Compilation, 1989.* Washington, DC: Tobacco Institute; 1990;24.
13. US Bureau of the Census. *Statistical Abstract of the United States: 1990.* 110th ed. Washington, DC: US Bureau of the Census; 1990.
14. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986;42:121–130.
15. Laird NM, Ware JH. Random effects models for longitudinal data. *Biometrics.* 1982;38:963–974.
16. De Boor C. *A Practical Guide to Splines.* New York, NY: Springer-Verlag; 1987.
17. Advisory Commission on Intergovernmental Relations. *Cigarette Tax Evasion: A Second Look.* Washington, DC: Advisory Commission on Intergovernmental Relations; 1985.

Mass Media–Led Antismoking Campaign Can Remove the Education Gap in Quitting Behavior

Petra Macaskill, MAppStat, John P. Pierce, PhD, Judy M. Simpson, PhD, and David M. Lyle, MBBS, PhD

ABSTRACT

This study investigated whether the effective mass media-led antismoking campaigns in Australia had the traditional differential effect across educational levels. Our population surveys included random samples of 12 851 people before the campaign and 11 609 several years after the campaign had started. No statistically significant differences were found in quitting across education levels in three of the four subgroups. Mass media-led antismoking campaigns may play an important role in getting the antismoking message to the less educated. (*Am J Public Health.* 1992;82:96–98)

Introduction

A major public health concern in developed countries is the increasing gap in smoking prevalence between educational groups.[1] In the United States, Canada, and Norway, smoking in the least educated groups is approximately twice as prevalent as in the most educated groups, and the rate of decline in smoking behavior is three to nine times lower.[2] Evidence indicates that those who are less educated obtain information from visual media (such as television) more often than those who are more educated.[3] If one barrier to less educated people changing their smoking behavior is the lack of access to motivational information,[4] use of a mass media–led antismoking campaign may reduce smoking prevalence overall without widening the gap between educational groups.

During the 1980s, mass media–led antismoking campaigns were conducted in two Australian capital cities, Sydney and Melbourne.[5,6] The Sydney campaign began in mid-1983, 1 year before the start of the Melbourne campaign. Using Melbourne as a control, Sydney showed a marked decline of 2.5% in smoking prevalence, a decline attributable to the campaign in the first year.[7] Subsequent results showed a similar decline in Melbourne in its first campaign year, and a continued decline over the next 4 years for men in each city.[8] Among adults, the decline in prevalence was primarily owed to quitting.[7]

During the years in which such campaigns were run, more than two thirds of the communities in both cities recalled television advertisements depicting the health consequences of smoking.[8] In this paper, we investigate whether the effects of this campaign differed by educational level.

Methods

A subset of the large Australian database was used for this study, which followed a "before-after" design.[8] Members of the "before" group (n = 12 851), aged 25 to 54, were interviewed in Sydney and Melbourne over a year starting in mid-

Petra Macaskill, Judy M. Simpson, and David M. Lyle are with the Department of Public Health, University of Sydney, in New South Wales, Australia. John P. Pierce is with the Cancer Center and Department of Community and Family Medicine of the University of California-San Diego in La Jolla.

Requests for reprints should be sent to John P. Pierce, PhD, Cancer Center, University of California-San Diego, 0901, La Jolla, CA 92093.

This paper was submitted to the journal August 30, 1990, and accepted with revisions July 18, 1991.

TABLE 1—Sample Sizes for Men and Women by Educational Level, %

	Sydney				Melbourne			
	Men		Women		Men		Women	
	1983	1988	1983	1988	1983	1988	1983	1988
Total sample	(3449)	(2973)	(3507)	(3132)	(2797)	(2522)	(3098)	(2982)
Age group, y								
25–34	43	44	44	45	45	49	44	49
35–44	36	34	33	33	34	31	33	30
45–54	22	22	23	23	21	20	23	21
Level of education								
Up to 9 years	21	19	27	24	28	23	37	33
Intermediate high school	18	16	29	28	12	13	17	17
Completed high school	27	30	22	22	25	27	23	24
Some university	34	35	22	26	35	37	23	26

1982. Members of the "after" group (n = 11 609), aged 30 to 59 to control for a possible birth cohort effect, were interviewed 5 years later. The sampling method for the Australian study has been described elsewhere.[6–8] Smokers were defined as those who said they smoked tobacco products. In two separate subsamples, self-reported smoking was validated with saliva analysis for cotinine.[9]

We divided subjects according to age in 1982/1983: 25 to 35, 35 to 44, and 45 to 54 years old. Level of education was categorized into four groups: up to 9 years of schooling, a 10-year certificate, a 12-year certificate, and some college education.

We used the rate ratio to measure changes in smoking behavior because this method controls for differences in initial smoking prevalence between groups by expressing the "after" prevalence as a percentage of the "before" prevalence. The relative decline in prevalence is thus (100 − rate ratio)%. The Mantel-Haenszel method[10] was used to determine the age-adjusted rate ratio and its 95% confidence interval for both genders within each city. We used multiplicative (rate ratio) models[11,12] to assess the probability of being a smoker as a function of age, year of interview, and level of education, and we calculated the probabilities separately for men and women in both Sydney and Melbourne. Backward elimination of terms in the saturated model was used to test the significance of interactions and main effects. In particular, a statistically significant ($P < .05$) interaction between level of education and year of interview would indicate that the rate of change in smoking prevalence between the two surveys differed across educational groups.

TABLE 2—Smoking Rates and Rate Ratios for Sydney and Melbourne Men, by Educational Level

City	Level of Education	% Smokers 1983	% Smokers 1988	Age-adjusted Rate Ratio (%)	95% CI
Sydney	Up to 9 years	56	46	81	(71,92)
	Intermediate high school	49	39	80	(69,92)
	Completed high school	41	35	88	(80,97)
	University	33	24	78	(71,87)
Melbourne	Up to 9 years	53	42	79	(69,89)
	Intermediate high school	50	42	83	(71,97)
	Completed high school	47	38	83	(74,92)
	Some university	35	27	81	(73,89)

Note. CI = confidence interval.

TABLE 3—Smoking Rates and Rate Ratios for Sydney and Melbourne Women, by Educational Level

City	Level of Education	% Smokers 1983	% Smokers 1988	Age-adjusted Rate Ratio (%)	95% CI
Sydney	Up to 9 years	41	34	84	(75,95)
	Intermediate high school	35	31	92	(83,102)
	Completed high school	31	24	81	(72,92)
	University	24	20	88	(78,99)
Melbourne	Up to 9 years	37	35	96	(87,106)
	Intermediate high school	33	30	92	(80,106)
	Completed high school	32	29	92	(82,103)
	Some university	29	20	77	(68,87)

Note. CI = confidence interval.

Results

The distributions of age and of educational level were similar for the 1982/1983 and 1987/1988 groups (Table 1), but adjustment for those variables was still necessary because of their strong association with smoking prevalence.

Tables 2 and 3 show smoking prevalence and rate ratios for men and women, respectively, indicating a clear gradient for smoking prevalence across educational levels: smoking prevalence in the least educated group is generally 70% greater than it is in the most educated group. The rate ratios show an overall de-

cline in prevalence as they are consistently below 100%. For both Sydney and Melbourne men and for Sydney women, the rate ratios for the four levels of education are not significantly different; i.e., after controlling for age, the relative decrease in smoking of the least educated group is not significantly different from that of the most educated group. The pattern differs, however, for Melbourne women, for whom there is a significant difference between the age-adjusted rate ratios of the four educational levels. The rate ratios vary from 96% (least educated group) to 77% (most educated group); the 23% decline in prevalence in the most educated group is the greatest seen among any group in either city, and the 5% to 9% declines among the other educational groups of Melbourne women are the lowest rates of decline of all groups studied.

Discussion

In Sydney and Melbourne, a significant decline in smoking prevalence occurred between the survey years 1982/1983 and 1987/1988.[7,8] This decline contrasts with the relatively stable smoking levels of the previous decade,[8,13] and the drop has been attributed to the mass media–led antismoking campaigns.[7,8] Despite the large sample sizes, we found no evidence that the gap in smoking prevalence between the educational groups increased over the study period in three of the four study groups. The exception occurred among Melbourne women, where only the higher educated showed a decline in smoking. The same group, Melbourne women, did not show any long-term decline in smoking clearly attributable to the campaign.[8]

These results are among the first reported examples of a health promotion–motivated behavioral change that did not lead to an increase in the gap between educational levels. They suggest that the much larger decreases in smoking prevalence seen among better-educated groups in previous studies may have been related to differences in exposure to motivational material rather than to differences in behavior-changing skills. To address the needs of smokers of all educational levels, the public health movement should conduct antismoking campaigns using the full powers of visual mass media. □

Acknowledgments

The data used in this study were collected as part of the evaluation for both the "Quit for Life" program conducted in Sydney and the Quit campaign conducted in Melbourne. We wish to thank Bev Higham of "Quit for Life" and the directorate of the drug offensive for their support and partial funding for the project. Additional funding for data collection came from the Anticancer Council of Victoria as part of their evaluation of the Melbourne Quit campaign.

References

1. Pierce JP, Fiore MC, Novotny TE, Hatziandreu E, Davis R. Trends in cigarette consumption in the United States: educational differences are increasing. *JAMA*. 1989;261:56–60.
2. Pierce JP. International trends in cigarette smoking. *Am J Public Health*. 1989;79:1–6.
3. Roberts DF, Maccoby N. Effects of mass communication. In: Lindzey G, Aronson E, eds. *Handbook of Social Psychology*. 3rd ed. Reading, Mass: Addison Wesley; 1984;2:539–598.
4. Rogers EM, Shoemaker F. *Communication of Innovations: A Cross-Cultural Approach*. 2nd ed. New York, NY: The Free Press; 1971.
5. Pierce JP, Dwyer T, Chamberlain A, et al. Targeting the smoker in an antismoking campaign. *Prev Med*. 1987;16:816–824.
6. Pierce JP, Dwyer T, Frape G, et al. Evaluation of the Sydney "quit for life" antismoking campaign: part I. Achievement of intermediate goals. *Med J Aust*. 1986;144:341–344.
7. Dwyer T, Pierce JP, Hannan CD, et al. Evaluation of Sydney "quit for life" antismoking campaign: part II. Changes in smoking prevalence. *Med J Aust*. 1986;144:344–347.
8. Pierce JP, Macaskill P, Hill DJ. Long-term effectiveness of mass media anti-smoking campaigns in Australia. *Am J Public Health*. 1990;80:565–569.
9. Pierce JP, Dwyer T, DeGiusto E, et al. Cotinine validation of self-reported smoking in commercially run community surveys. *J Chronic Dis*. 1987;40:689–695.
10. Rothman KR. *Modern Epidemiology*. Boston, Mass: Little Brown and Co; 1986.
11. Wacholder S. Binomial regression in GLIM: estimating risk ratios and risk differences. *Am J Epidemiol*. 1986;123:174–179.
12. Payne CD, ed. *The GLIM System Release 3.77 Manual*. Oxford, England: Numerical Algorithms Group; 1987.
13. Hill DJ, White VM, Gray NJ. Measures of tobacco smoking in Australia, 1974–1986, by means of a standard method. *Med J Aust*. 1988;149:10–12.

Demographic and Socioeconomic Differences in Beliefs about the Health Effects of Smoking

Ross C. Brownson, PhD, Jeannette Jackson-Thompson, PhD, MSPH, Joan C. Wilkerson, BA, James R. Davis, PhD, Nancy W. Owens, MEd, and Edwin B. Fisher, Jr, PhD

ABSTRACT

To assess sociodemographic differences in beliefs about the health effects of cigarette smoking and passive smoke exposure, we recently surveyed 2092 adults in St. Louis and Kansas City, Mo. The percentages of respondents who knew that smoking causes lung cancer, emphysema, and heart disease were 76.7, 74.1, and 67.2, respectively. After multivariate adjustment, knowledge about smoking's health effects was generally lower among women, older respondents, those of lower education level, and current smokers. Blacks were generally less likely to appreciate the health effects of active smoking, but were more likely to acknowledge the health effects of passive smoking. (*Am J Public Health*. 1992;82:99–103)

Introduction

The harmful health effects of cigarette smoking are well documented. Cigarette smoking was first linked to lung cancer in 1950,[1] and in 1964 the US surgeon general found that smoking was causally associated with lung and laryngeal cancers and chronic bronchitis.[2] Cigarette smoking is now considered the leading avoidable cause of mortality in the United States, accounting for approximately 434 000 deaths in 1988.[3]

The health hazards of passive smoke exposure have only recently been established. Passive smoking is now regarded as a cause of lung cancer in healthy nonsmokers,[4,5] accounting for approximately 3800 lung cancer deaths per year in the United States.[3] In addition, the children of parents who smoke have a higher frequency of respiratory infections and decreased lung function as the lungs mature.[4,5]

Public knowledge and beliefs about the health risks of smoking and passive smoke exposure have increased substantially over the past several decades. For example, between 1964 and 1987, the percentage of adults who knew that smoking causes lung cancer increased from 66 to nearly 90.[6,7] It also appears that a majority of the population is aware of the health hazards of passive smoking. A nationwide study in 1986 found that 81% of respondents knew that passive smoke exposure was hazardous to nonsmokers' health.[8]

Despite considerable gains in public knowledge about smoking, substantial numbers of smokers are still unaware of or do not accept the health risks of smoking.[4] In addition, few studies have examined related issues such as knowledge of nicotine addiction, the health benefits of quitting, and passive smoking health issues. Even fewer studies have evaluated knowledge about smoking and passive smoking among high-risk groups such as minorities and persons of lower socioeconomic status.[9–11]

This paper presents information on knowledge and beliefs about cigarette smoking and passive smoke exposure that was collected in a recent survey of an urban adult population.

Methods

The present study was conducted to provide planning and evaluation data for a smoking prevention and cessation intervention in a high-risk, inner-city population. From March through September 1990, computer-assisted telephone interviews were conducted with 2092 adults 18 years of age and older. Study subjects were randomly selected from among residents of approximately 60 census tracts in St. Louis and Kansas City, Mo. Techniques for the sample selection and survey were based on the Behavioral Risk Factor Surveillance System.[12] A two-stage, random-digit dialing technique was used to collect data from residents of these lower-socioeconomic-status, predominantly Black neighborhoods. During the first stage, most nonworking and ineligible telephone numbers were eliminated. In the second stage, a respondent was identified from among all adult residents of the household.[12] The response rate among eligible households was high (93%). Table 1

Ross C. Brownson is with the Division of Chronic Disease Prevention and Health Promotion; Jeannette Jackson-Thompson and James R. Davis are with the Bureau of Smoking, Tobacco, and Cancer; and Joan C. Wilkerson is with the Missouri Cancer Registry—all at the Missouri Department of Health, Columbia. Nancy W. Owens is with the Grace Hill Neighborhood Health Center, Inc., St. Louis. Edwin B. Fisher is with the Center for Health Behavior Research, Washington University School of Medicine, St. Louis.

Requests for reprints should be sent to Ross C. Brownson, PhD, Director, Division of Chronic Disease Prevention and Health Promotion, Missouri Department of Health, 201 Business Loop 70 West, Columbia, MO 65203.

This paper was submitted to the journal January 28, 1991, and accepted with revisions July 18, 1991.

TABLE 1—Respondent Characteristics by Sociodemographic Grouping and Smoking Status, St. Louis and Kansas City, 1990

Characteristic	No.	%
Age (y)		
18–34	767	36.7
35–54	612	29.2
55+	679	32.5
Unknown/refused	34	1.6
Gender		
Male	725	34.7
Female	1367	65.3
Race		
White	434	20.7
Black	1564	74.8
Other	75	3.6
Unknown/refused	19	0.9
Education level		
College graduate	249	11.9
Some college	507	24.2
High school graduate	743	35.5
Less than high school graduate	572	27.3
Unknown/refused	21	1.0
Smoking status		
Never smoker	1009	48.2
Former smoker	406	19.4
Current smoker	677	32.4

presents the distribution of the study population by selected factors.

The study questionnaire was a modified version of the Behavioral Risk Factor Surveillance System instrument.[12] (A copy of the questionnaire is available from the first author on request.) Tobacco-related questions focused on smoking history, knowledge and beliefs about the health effects of smoking and passive smoke exposure, and attitudes toward tobacco use and tobacco control policies. Questions on smoking history and knowledge about health effects of smoking were based on standardized questions from the American Lung Association,[13] the Adult Use of Tobacco Survey,[8] and National Cancer Institute-funded researchers.[14,15]

Initially, simple frequencies were calculated by sociodemographic grouping.[16,17] Also, logistic regression was used to calculate odds ratios (ORs) and 95% confidence intervals, while simultaneously adjusting for other independent variables.[18] Trends in proportions were also evaluated.[19,20]

Results

Table 2 presents respondent knowledge and beliefs about the health effects of cigarette smoking. The vast majority of respondents (88.4%) believed that smoking is harmful to health. Belief in harmful health effects was inversely related to age and positively correlated with education level ($P < .001$). Current smokers were significantly less likely than never smokers to acknowledge the health effects of smoking (OR = 0.5). The overall percentages of respondents who knew smoking causes specific diseases such as lung cancer, emphysema, and heart disease were 76.7, 74.1, and 67.2, respectively (not shown in a table). As with the general question regarding the harmful health outcomes of smoking, knowledge of smoking as a cause of these three diseases was generally lower among older respondents, women, less-educated subjects, and current smokers (Table 2).

Sociodemographic differences regarding beliefs about the addictive nature of cigarette smoking and the health benefits of quitting are shown in Table 3. The majority of respondents knew that cigarette smoking was physically addicting (88.6%), and that there are health benefits from quitting smoking (83.2%). After multivariate adjustment, respondents 55 years of age and older (OR = 0.6) and current smokers (OR = 0.7) were significantly less aware of the addictive nature of smoking. Less-educated respondents were less likely to acknowledge the health benefits of quitting.

Table 4 shows the distribution of the study population regarding beliefs about passive smoke exposure. A majority of respondents believed passive smoke exposure to be harmful to nonsmokers' health (78.2%), harmful to young children (87.2%), and annoying (65.5%). Results regarding health effects and annoyance of passive smoking were inversely correlated with age ($P < .001$). Blacks were more likely than Whites to agree that passive smoke exposure is harmful to young children's health (OR = 1.3) and to find passive smoke annoying (OR = 1.2). Current smokers were significantly less likely than never smokers to acknowledge the effects of passive smoking on nonsmokers' health (OR = 0.3) and children's health (OR = 0.2) and were less often annoyed by passive smoke (OR = 0.1).

Discussion

Since relatively few studies have described sociodemographic differences in beliefs about the health effects of smoking and passive smoke exposure, our results provide some useful insights. Although differences were not always statistically significant after multiple adjustment, we found knowledge about the health effects of smoking and passive smoking to be lower among older age groups, women, less-educated respondents, and current smokers. Racial differences were interesting in that Blacks were generally less likely to appreciate the health effects of active smoking but were more likely to acknowledge passive smoking's health effects.

The percentages of respondents in our study who knew that smoking causes lung cancer, emphysema, and heart disease were slightly lower than those found in two recent national studies.[7,8] For example, the American Lung Association[7] found that a majority of respondents knew smoking is a cause of lung cancer (87%) and emphysema (85%). The corresponding figures for our study were 77% and 74%, respectively. As in previous studies,[7,8] subjects in our study were most likely to recognize lung cancer as a consequence of smoking and least likely to recognize heart disease.

Numerous perspectives in behavioral science (e.g., the health belief model) assert that an individual needs to feel susceptible to a health risk prior to changing behavior.[4,21] Our data show that for both general health effects and specific disease threats, smokers are significantly less likely than nonsmokers or former smokers to acknowledge the harmful effects of smoking. For example, 33% of current smokers in our study, almost twice as many as among never smokers, were unaware that smoking is a cause of lung cancer. More striking, perhaps, are comparisons for heart disease. Smoking is the

TABLE 2—Beliefs about the Health Effects of Cigarette Smoking by Sociodemographic Grouping and Smoking Status, St. Louis and Kansas City, 1990

Characteristic	Smoking Is Harmful to Health			Smoking Is Cause of: Lung Cancer			Emphysema			Heart Disease		
	%	OR[a]	95% CI	%	OR[a]	95% CI	%	OR[a]	95% CI	%	OR[a]	95% CI
Age (y)												
18–34	95.6	1.0		86.2	1.0		73.7	1.0		71.8	1.0	
35–54	87.3	0.4	0.3,0.6	74.5	0.5	0.4,0.7	78.4	1.4	1.1,1.8	68.4	0.9	0.7,1.1
55+	81.7	0.2	0.1,0.3	68.3	0.4	0.3,0.5	71.2	1.0	0.8,1.3	61.8	0.7	0.6,0.9
Gender												
Male	89.4	1.0		80.7	1.0		76.1	1.0		71.7	1.0	
Female	87.9	0.8	0.6,1.1	74.6	0.6	0.5,0.8	73.0	0.9	0.7,1.1	64.9	0.7	0.6,0.9
Race[b]												
White	89.6	1.0		80.0	1.0		79.9	1.0		70.9	1.0	
Black	88.3	0.9	0.7,1.3	76.0	0.9	0.7,1.1	73.1	0.9	0.7,1.1	66.7	0.9	0.8,1.2
Education level												
College graduate	95.2	1.0		90.0	1.0		91.1	1.0		79.8	1.0	
Some college	89.9	0.6	0.3,1.1	81.1	0.6	0.4,1.0	78.1	0.5	0.3,0.8	71.0	0.8	0.5,1.1
High school graduate	88.7	0.6	0.3,1.0	76.7	0.5	0.3,0.8	71.7	0.4	0.2,0.6	66.8	0.6	0.5,0.9
Less than high school graduate	84.3	0.5	0.3,1.0	67.7	0.4	0.3,0.6	67.0	0.3	0.2,0.5	60.1	0.5	0.4,0.8
Smoking status												
Never smoker	91.1	1.0		83.0	1.0		74.9	1.0		67.5	1.0	
Former smoker	91.6	1.1	0.8,2.0	78.1	0.7	0.6,1.0	80.7	1.3	1.0,1.8	71.4	1.2	0.9,1.6
Current smoker	82.9	0.5	0.3,0.6	66.9	0.4	0.3,0.5	69.2	0.7	0.6,0.9	64.5	0.9	0.7,1.1

Note. OR = odds ratio; CI = confidence interval.
[a]Each OR was adjusted by logistic regression for all other variables in the table.[18]
[b]Results for other racial/ethnic groups are not shown because of small numbers.

cause of larger numbers of heart disease than lung cancer deaths,[3,4] yet almost 40% of current smokers, as opposed to 29% of former smokers, indicated a lack of awareness of this grave risk. This reaffirms that many smokers underestimate or deny the serious health risks of smoking.[4]

Although the 1988 surgeon general's report clearly determined that cigarette smoking is a potent addiction,[22] few data are available on public understanding of this addiction. A 1986 survey in Canada found that 76.5% of respondents believed cigarette smoking to be addicting.[23] Of the subjects we surveyed, 89% felt that cigarette smoking was addicting. Knowledge of this addiction was significantly lower among persons 55 years of age and older and among current smokers.

Understanding of the health effects of passive smoking was relatively high in our study. Overall, 78% of those surveyed indicated that passive smoke exposure was harmful to nonsmokers' health. This compares with 81% of those surveyed in two recent national studies.[8,24] The majority of respondents acknowledged the detrimental effects of passive smoke exposure on children. Among subgroups, knowledge was lower in older age groups, women, Whites, and former or current smokers. Most of those surveyed were annoyed by passive smoke exposure, as reported earlier.[25] Data on passive smoking annoyance affirm the declining social acceptability of smoking in public places. Nearly

TABLE 3—Beliefs about Addiction to Cigarette Smoking and the Benefits of Quitting by Sociodemographic Grouping and Smoking Status, St. Louis and Kansas City, 1990

Characteristic	Smoking Is Addicting[a]			Health Benefits from Quitting[a]		
	%	OR[b]	95% CI	%	OR[b]	95% CI
Age (y)						
18–34	90.9	1.0		82.9	1.0	
35–54	90.7	1.1	0.8,1.6	87.1	1.5	1.1,2.0
55+	84.8	0.6	0.5,0.9	80.9	1.0	0.8,1.3
Gender						
Male	89.2	1.0		83.6	1.0	
Female	88.4	0.9	0.7,1.2	83.0	1.0	0.8,1.2
Race[c]						
White	90.3	1.0		88.7	1.0	
Black	88.5	0.9	0.6,1.2	82.3	0.8	0.6,1.1
Education level						
College graduate	91.6	1.0		93.2	1.0	
Some college	90.7	1.1	0.7,1.8	84.6	0.5	0.3,0.9
High school graduate	89.6	1.0	0.6,1.6	82.0	0.4	0.3,0.7
Less than high school graduate	84.8	0.8	0.5,1.3	79.6	0.4	0.2,0.7
Smoking status						
Never smoker	89.7	1.0		83.8	1.0	
Former smoker	91.1	1.2	0.8,1.9	85.5	1.1	0.8,1.6
Current smoker	86.0	0.7	0.5,0.9	81.2	0.8	0.6,1.1

Note. OR = odds ratio; CI = confidence interval.
[a]Includes respondents who "strongly agree" or "agree."
[b]Each OR was adjusted by logistic regression for all other variables in the table.[18]
[c]Results for other racial/ethnic groups are not shown because of small numbers.

TABLE 4—Beliefs about the Health Effects and Annoyance of Passive Smoke Exposure by Sociodemographic Grouping and Smoking Status, St. Louis and Kansas City, 1990

Characteristic	Passive Smoke Exposure Is								
	Harmful to Nonsmokers' Health			Harmful to Young Children's Health[a]			Annoying[b]		
	%	OR[c]	95% CI	%	OR[c]	95% CI	%	OR[c]	95% CI
Age (y)									
18–34	86.2	1.0		93.7	1.0		77.7	1.0	
35–54	77.2	0.6	0.5,0.9	85.5	0.4	0.3,0.6	61.8	0.5	0.4,0.7
55+	70.5	0.4	0.3,0.5	81.6	0.3	0.2,0.4	55.2	0.3	0.2,0.4
Gender									
Male	78.3	1.0		88.3	1.0		60.1	1.0	
Female	78.4	0.9	0.7,1.1	86.7	0.7	0.5,0.9	68.3	1.2	1.0,1.5
Race[d]									
White	77.7	1.0		84.3	1.0		61.1	1.0	
Black	78.3	1.0	0.7,1.3	87.9	1.3	1.0,1.8	66.6	1.2	1.0,1.5
Education level									
College graduate	85.9	1.0		92.8	1.0		75.5	1.0	
Some college	79.9	0.8	0.5,1.2	86.2	0.6	0.3,1.0	70.6	0.9	0.6,1.4
High school graduate	79.4	0.8	0.5,1.2	88.7	0.8	0.5,1.3	66.1	0.8	0.6,1.2
Less than high school graduate	72.6	0.7	0.5,1.1	84.1	0.7	0.4,1.2	56.1	0.7	0.5,1.0
Smoking status									
Never smoker	86.3	1.0		93.5	1.0		81.5	1.0	
Former smoker	80.8	0.8	0.6,1.1	86.7	0.5	0.3,0.8	71.9	0.7	0.6,1.0
Current smoker	65.4	0.3	0.2,0.4	78.4	0.2	0.2,0.3	38.1	0.1	0.1,0.2

Note. OR = odds ratio; CI = confidence interval.
[a]Includes respondents who "strongly agree" or "agree."
[b]Includes respondents who found passive smoke "very annoying" or "somewhat annoying."
[c]Each OR was adjusted by logistic regression for all other variables in the table.[18]
[d]Results for other racial/ethnic groups are not shown because of small numbers.

40% of the current smokers we surveyed were annoyed by others' cigarette smoke.

Our findings may not be entirely representative of the general population, since we surveyed a low-income population. Thirty percent of the respondents in our study had household incomes of less than $10 000 per year, which compares with 22% of the overall state population.[26] However, since low socioeconomic status is a strong predictor of smoking prevalence,[9] the need for additional research in low-income populations is apparent.

The current study demonstrates some important sociodemographic differences in knowledge and beliefs about the health effects of cigarette smoking and passive smoke exposure. The patterns that we identified should be considered when targeting public education and advocacy efforts to the segments of the population at highest risk for smoking and passive smoke exposure. □

Acknowledgments

This study was supported in part by National Heart, Lung, and Blood Institute Grant 5 R01 HL43555. We are grateful to supervisors and surveyors from Manpower, Inc., who conducted data collection.

References

1. Wynder EL. Tobacco and health: a review of the history and suggestions for public policy. *Public Health Rep.* 1980;103:8–18.
2. US Department of Health, Education, and Welfare. *Smoking and Health. Report of the Advisory Committee to the Surgeon General of the Public Health Service.* Washington, DC: US Government Printing Office; 1964; PHS publication no. 1103.
3. Centers for Disease Control. Smoking-attributable mortality and years of potential life lost—United States, 1988. *MMWR.* 1991;40:62–71.
4. US Department of Health and Human Services. *Reducing the Health Consequences of Smoking. 25 Years of Progress: A Report of the Surgeon General 1989.* Rockville, Md: Office on Smoking and Health; 1989; DHHS publication no. (CDC) 89-8411.
5. US Department of Health and Human Services. *The Health Consequences of Involuntary Smoking. A Report of the Surgeon General.* Washington, DC: US Government Printing Office; 1986; DHHS publication no. (CDC) 87-8398.
6. US Department of Health, Education, and Welfare. *Use of Tobacco. Practices, Attitudes, Knowledge, and Beliefs. United States—Fall 1964 and Spring 1966.* Washington, DC: National Clearinghouse for Smoking and Health; 1969.
7. American Lung Association. Smokers know risks but don't say "no" to cigarettes. [Press release]. New York: American Lung Association; December 4, 1987.
8. US Department of Health and Human Services. *Report of the 1986 Adult Use of Tobacco Survey.* Rockville, Md: Office on Smoking and Health; 1986.
9. Novotny TE, Warner KE, Kendrick JS, Remington PL. Smoking by Blacks and Whites: socioeconomic and demographic differences. *Am J Public Health.* 1988;78:1187–1189.
10. Fiore MC, Novotny TE, Pierce JP, Hatziandreu EJ, Patel KM, Davis RM. Trends in cigarette smoking in the United States. The changing influence of gender and race. *JAMA.* 1989;261:49–55.
11. Pierce JP, Fiore MC, Novotny TE, Hatziandreu EJ, Davis RM. Trends in cigarette smoking in the United States. Educational differences are increasing. *JAMA.* 1989;261:56–60.
12. Remington PL, Smith MY, Williamson DF, Anda RF, Gentry EM, Hogelin GC. Design, characteristics, and usefulness of state-based behavioral risk factor surveillance: 1981–1987. *Public Health Rep.* 1988;103:366–375.
13. American Lung Association. Summary of results of the 1989 survey on public attitudes toward smoking conducted by the Gallup Organization, New York: American Lung Association; November, 1989.
14. Brownson RC, DiLorenzo TM, Van Tuinen M, Finger WW. Patterns of cigarette and smokeless tobacco use among children and adolescents. *Prev Med.* 1990;19:170–180.

15. Mattson ME, Cummings KM, Lynn WR, Giffen C, Corle D, Pechacek T. Evaluation plan for the community intervention trial for smoking cessation (COMMIT). *Int J Community Health Educ.* In press.
16. Statistical Analysis System Institute Inc. *SAS User's Guide: Basics, Version 5 Edition.* Cary, NC: SAS Institute; 1985.
17. Statistical Analysis System Institute Inc. *SAS User's Guide: Statistics, Version 5 Edition.* Cary, NC: SAS Institute; 1985.
18. Engelman L. Stepwise logistic regression. In: Dixon WJ, ed. *BMDP Statistical Software.* Berkeley, Calif: University of California Press; 1988;941–969.
19. Mantel N. Chi-square tests with one degree of freedom: extensions of the Mantel-Haenszel procedure. *J Am Stat Assoc.* 1963;58:690–700.
20. Rothman KJ, Boice JD Jr. *Epidemiologic Analysis with a Programmable Calculator.* Boston, Mass: Epidemiology Resources; 1982.
21. Rosenstock IM. The health belief model and preventive health behavior. *Health Educ Monogr.* 1974;2:354–386.
22. US Department of Health and Human Services. *The Health Consequences of Smoking. Nicotine Addiction: A Report of the Surgeon General.* Rockville, Md: Office on Smoking and Health; 1988; DHHS publication no. (CDC) 88-8406.
23. Gallup. Canadian Gallup Poll. October, 1986.
24. American Cancer Society. *1987 Survey of Public Awareness and Use of Cancer Tests.* New York, NY: American Cancer Society; 1988.
25. Centers for Disease Control. Passive smoking: beliefs, attitudes, and exposures—United States, 1986. *MMWR.* 1988;37:239–241.
26. US Department of Commerce. *1980 Census of Population. Detailed Population Characteristics. Missouri.* Washington, DC: US Government Printing Office; 1983; No. PC80-1-D27.

Geographic and Gender Variations in Total Tobacco Use

Donald R. Shopland, Steve J. Niemcryk, PhD, and Katherine M. Marconi, PhD

ABSTRACT

This study is the first to provide complete information on prevalence rates by gender and geographic variation for each type of tobacco product used in the United States. Results indicate that, in nearly half of all states, total tobacco use in men exceeded 40% and, in four states, exceeded 50%. In women, only Nevada, Kentucky, and Michigan reported prevalence exceeding 33%. Results also indicate, however, that concurrent use of multiple tobacco forms is relatively rare.

Substantial regional variation in male total tobacco use was evident, with southern males exhibiting the highest prevalence rate (44.6%). As a result, this region represents a target group in special need of comprehensive and effective tobacco use interventions. (*Am J Public Health.* 1992;82:103–106)

Introduction

National surveys conducted since 1955 have systematically documented the major demographic and socioeconomic variables associated with tobacco use, especially cigarettes.[1–4] Considerably less attention has been paid to differences in prevalence by geographic location, and until 1985, no single national study existed that was capable of providing complete prevalence estimates for each type of tobacco product used (i.e., cigarettes, pipes, cigars, chewing tobacco, and snuff) within various geographic strata.

In September 1985, the Office on Smoking and Health, DHHS, added a brief series of questions on current tobacco use practices to the Current Population Survey (CPS). Owing to its large sample size and unique design, the September CPS is the only survey currently capable of producing reliable prevalence estimates for individual tobacco products by geographic location, including each of the 50 states and the District of Columbia.[5,6] The purpose of this analysis is to use the 1985 CPS to provide estimates of the prevalence of use of individual tobacco products, by gender, and of the use of at least one tobacco product for each region, division, and state.

Methods

The methodology of the CPS has been described in detail elsewhere.[7] Briefly, based on the 1980 census file and on a system of stratified random sampling, eligible households from each state and the District of Columbia were identified. For each eligible household, questionnaires were completed for each household member aged 16 or older (n = 114 342). Proxy respondents were used to provide data on approximately 45% of the sample.

For this analysis, only data from individuals aged 20 or older (n = 105 225) were used to reduce the potential problem of using proxy responses for younger age groups. Subjects who reported using tobacco in any form—cigarettes, cigars, pipes, snuff, or chewing tobacco—at the time of interview were classified as current users. Definitions used to determine tobacco use status in the CPS were identical to those used in the National Health Interview Survey (NHIS), as well as in other surveys conducted by various federal agencies.

Donald R. Shopland and Katherine M. Marconi are with the Cancer Control Sciences Program at the National Cancer Institute in Bethesda, Maryland; Steve J. Niemcryk is with R.O.W. Sciences, Inc., in Rockville, Maryland.

Requests for reprints should be sent to Donald R. Shopland, Coordinator, Smoking and Tobacco Control Program, National Cancer Institute, EPN-241, 9000 Rockville Pike, Bethesda, MD 20892.

This paper was submitted to the journal March 4, 1991, and accepted with revisions July 9, 1991.

TABLE 1—Percent Of Tobacco Use In Persons 20 Years of Age or Older, by Gender and Geographic Location

	Tobacco Use in Adult Males								Tobacco Use in Adult Females			
	None	Cigarette	Snuff	Chewing Tobacco	Cigar	Pipe	At Least One Tobacco Product	CI[a]	None	Cigarette	At Least One Tobacco Product	CI[a]
United States (105 225)[b]	60.32	32.66	1.84	3.95	2.14	2.40	39.67	0.78	73.41	25.81	26.59	0.72
Northeast (26 186)	63.12	31.00	0.94	1.48	3.15	2.57	36.89	1.84	74.00	25.77	26.01	1.56
New England (9434)	64.34	30.25	0.42	0.74	3.67	2.79	35.66	3.64	72.25	27.58	27.75	3.17
Middle Atlantic (16 752)	62.71	31.26	1.11	1.73	2.98	2.49	37.30	2.14	74.59	25.16	25.40	1.79
North Central (26 154)	60.45	32.20	2.07	3.43	2.29	3.01	39.55	1.71	72.54	27.33	27.46	1.48
East North Central (16 388)	60.07	32.83	1.70	2.90	2.53	3.20	39.94	2.04	71.31	28.55	28.70	1.80
West North Central (9766)	61.34	30.70	2.95	4.69	1.72	2.54	38.66	3.11	75.40	24.51	24.60	2.60
South (31 646)	55.44	36.04	2.48	6.22	1.89	2.17	44.56	1.50	72.11	26.03	27.89	1.22
South Atlantic (18 014)	56.53	35.89	1.64	5.35	2.12	2.23	43.46	2.11	71.79	26.19	28.20	1.71
East South Central (4985)	51.49	37.33	2.57	9.58	1.65	2.12	48.51	3.52	71.77	25.95	28.22	2.80
West South Central (8647)	55.99	35.54	3.82	5.65	1.66	2.11	44.01	2.69	72.84	25.82	27.16	2.20
West (21 239)	65.57	29.19	1.41	3.34	1.29	1.82	34.43	1.89	76.28	23.47	23.72	1.62
Mountain (8854)	62.20	29.92	2.36	5.61	1.46	2.04	37.81	3.72	75.44	24.27	24.57	3.17
Pacific (12 385)	66.81	28.92	1.05	2.51	1.23	1.73	33.18	2.19	76.58	23.18	23.42	1.88

Note. Percentages are based on the number of male and female subjects within the individual geographic catagories.
[a]Confidence intervals are calculated for the rate of use of at least one tobacco product. [b] Numbers in parentheses indicate the number males and females in the present analysis.

Results

Approximately one third (32.8%) of adults aged 20 and older used at least one type of tobacco product in 1985. Nearly 40% of all men reported using at least one form of tobacco in 1985 compared with 26.6% of all women. (Table 1). Use of multiple forms of tobacco, however, was relatively rare among males and was almost nonexistent among females.

Among males, the prevalence of using specific tobacco products varied substantially by region (Table 1), with the South reporting the highest overall prevalence of male tobacco users (44.6%) and the West reporting the lowest (34.4%). Within regions, however, the percentage of females who used at least one product was only slightly higher than the prevalence of cigarette smoking alone whereas among males this difference varied considerably.

The prevalence of any tobacco use for males and females by individual states is presented in Tables 2 and 3. Rates are arranged in quintiles in descending order of prevalence for at least one form of tobacco use. For men, 10 of the 11 states with the highest prevalence of any tobacco use were southern states, and 9 of those 10 states were among the states with the highest prevalence of cigarette use. In nearly half of all states, the prevalence among men of using one or more tobacco product reached or exceeded 40%, and 12 states exhibited rates greater than 45% (Table 2).

For females, in only six states was the prevalence of using any tobacco product greater than 3% over that of using cigarettes alone (Table 3), and all six states are in the South. These states were the exception, however, because the prevalence of any tobacco use among females in all other states was less than 2 percent higher than that for cigarette use.

Discussion

Consistent with previously published reports,[4-8] the findings from the present study indicate that few females used tobacco in forms other than cigarettes and that rates of cigarette smoking among females did not vary significantly by region. Thus, the observed regional variation in total tobacco use is attributed almost exclusively to the patterns of male tobacco use, which was not limited to cigarettes. Levels of tobacco use among males in the different geographic areas differed substantially for total tobacco use, cigarettes, snuff, and chewing tobacco, but not for cigars and pipes.

The South clearly dominated other regions of the country in both cigarette use and use of at least one tobacco product among males. Given the substantially higher rates among southern men for both cigarettes and total tobacco use, southern females were expected to predominate in the rankings as well. Yet females from the South reported rates of cigarette use that were not substantially different from rates among women across the country overall, and the 10 states with the highest rates of cigarette smoking for females were about equally represented across the four regions. The absence of elevated smoking rates among southern women may indicate that the more traditional social norms and values historically associated with the South may have exerted a gender-specific effect on personal behavior that discouraged smoking by women but not by men. This would be analogous to contemporary norms regarding the social unacceptability of smokeless tobacco use by women or, the declining degree of public acceptance for smoking in various public settings.[9]

Prevalence estimates derived from the CPS are somewhat lower than those reported in other national surveys.[4] For example, in the 1985 NHIS, cigarette prevalence for women was 28.0% compared with 26.5% in the CPS (age-adjusted); the corresponding rates for men were 33.2% and 32.9% for NHIS and CPS, respectively.[10] Any differences observed between the CPS and NHIS can be at least partially attributed to use of proxy respondents. Nevertheless, based on this analysis, we estimate that approximately 54 million adult Americans used some form of tobacco in 1985.

Differences between men and women over the past 25 years in the documented rate of change in the prevalence of cigarette smoking has prompted some investigators to predict that cigarette

TABLE 2—Percent of Tobacco Use in Males 20 Years of Age and Older, by State in Descending Order of Prevalence of Overall Tobacco Use[a]

State	None	Cigarette	Snuff	Chewing Tobacco	Cigar	Pipe	At Least One Tobacco Product	CI[b]
West Virginia	43.33	37.73	10.57	13.72	1.45	3.57	56.66	7.31
Mississippi	45.66	38.24	5.90	11.61	2.61	2.83	54.34	7.56
Alaska	48.84	42.56	2.32	6.48	2.30	2.75	51.16	7.62
Kentucky	50.00	37.60	3.11	11.28	1.79	1.96	50.01	7.52
Arkansas	50.23	36.65	5.73	9.82	2.41	1.31	49.77	7.58
North Carolina	51.22	39.48	1.72	9.04	1.70	2.10	48.78	3.94
Georgia	52.93	38.86	0.74	8.22	2.05	1.05	47.06	7.12
Oklahoma	53.65	35.53	4.87	6.71	2.84	2.61	46.35	7.44
Tennessee	53.73	36.21	1.70	9.51	1.56	1.93	46.27	7.58
Alabama	53.75	37.84	1.08	6.88	1.05	2.07	46.24	8.46
Virginia	54.44	37.22	2.24	6.50	3.01	2.43	45.56	6.85
Kansas	54.56	34.25	3.29	8.54	1.76	3.57	45.44	7.39
Wyoming	55.13	31.64	3.64	12.07	0.42	2.29	44.87	8.30
Texas	56.67	35.40	3.81	4.74	1.50	2.05	43.33	3.83
New Mexico	56.88	32.32	5.28	5.50	1.68	1.62	43.12	7.05
Indiana	57.02	35.76	2.73	3.33	2.71	2.55	42.98	5.98
Nevada	57.15	38.26	1.41	3.06	0.46	1.41	42.85	8.58
Ohio	58.08	34.46	2.03	3.24	2.58	3.59	41.91	3.74
Louisiana	58.35	35.45	2.01	6.00	1.05	2.44	41.66	7.33
Delaware	58.55	35.17	0.67	2.62	1.97	3.88	41.45	8.12
Michigan	59.22	34.37	0.80	2.72	2.40	3.57	40.78	3.99
Arizona	59.42	33.60	1.99	3.96	2.26	2.30	40.58	8.22
Rhode Island	59.48	35.62	0.50	0.66	4.01	2.84	40.52	8.31
Vermont	59.85	35.31	0.68	4.29	1.08	4.88	40.14	7.93
Florida	60.10	34.17	1.15	1.73	2.33	2.37	39.90	3.61
New Hampshire	60.20	34.91	1.01	1.09	1.77	3.85	39.81	8.22
Pennsylvania	60.35	31.58	2.65	3.10	2.66	2.95	39.65	3.62
South Carolina	60.72	33.69	0.29	5.62	1.53	1.54	39.28	7.02
Minnesota	61.29	29.91	3.89	2.77	2.99	2.98	38.71	6.86
Colorado	61.41	30.44	1.33	6.95	1.27	2.82	38.59	7.78
North Dakota	61.94	28.41	6.24	5.62	1.04	1.54	38.06	6.69
Montana	62.08	25.77	5.18	8.63	1.23	2.18	37.91	6.99
South Dakota	62.29	29.89	2.02	5.74	1.00	1.68	37.71	6.49
District of Columbia	62.30	33.89	0.00	0.44	2.74	1.80	37.70	7.94
Iowa	62.40	32.56	1.80	4.69	0.48	1.41	37.60	7.28
Missouri	62.44	30.54	2.73	3.69	1.49	2.64	37.57	6.21
Illinois	62.49	31.23	1.12	2.43	2.55	2.58	37.51	3.81
New York	63.01	31.30	0.48	1.25	3.29	2.29	36.99	2.81
Maryland	63.49	31.19	0.42	1.98	1.78	2.89	36.51	6.35
Wisconsin	63.76	27.09	2.63	3.00	2.37	3.86	36.24	6.78
Connecticut	64.15	30.48	0.33	0.53	4.45	2.79	35.85	6.85
Maine	64.39	30.89	0.86	1.31	1.77	2.78	35.61	7.27
Oregon	64.61	26.60	2.96	5.20	1.56	2.51	35.40	7.93
Washington	64.78	29.33	1.94	5.63	1.07	2.20	35.22	7.80
Nebraska	65.42	26.15	1.57	6.70	1.73	2.30	34.58	6.88
New Jersey	65.77	30.64	0.10	0.65	2.78	2.22	34.23	3.60
Idaho	66.19	26.03	2.55	6.75	1.44	2.04	33.82	7.07
Massachusetts	66.50	28.09	0.24	0.34	4.15	2.39	33.51	3.82
California	67.68	28.81	0.70	1.68	1.20	1.59	32.32	2.83
Hawaii	68.47	29.88	0.24	0.47	1.34	0.71	31.53	7.06
Utah	76.94	18.88	0.86	3.15	1.21	0.62	23.07	8.30

[a]Crude rates. [b]Confidence intervals are computed for the rate of use of at least one tobacco product.

prevalence among women may exceed that among men by the middle of the 1990s.[1] Given the findings from this analysis, the high rate of total tobacco use by men also must be addressed. When all forms of tobacco use are considered, use of tobacco among males still far exceeds that among females. Moreover, certain regions of the country—particularly states in and around the South—will require more intensive efforts to curb smoking and tobacco use if the nation's health goals for the year 2000 are to be achieved.[1]

References

1. Haenszel W, Shimkin MB, Miller HP. Tobacco Smoking in the United States. Washington, DC: US Dept of Health, Education and Welfare; November 1956. Public Health Service Monograph 45.
2. Ahmed PI, Gleeson GA. Changes in cigar

TABLE 3—Percent of Tobacco Use in Females 20 Years of Age and Older, by State in Descending Order of Prevalence of Overall Tobacco Use[a]

State	None	Cigarette	At Least One Tobacco Product	CI[b]
Nevada	65.31	34.91	34.69[c]	7.92
Kentucky	65.69	33.41	34.31	6.83
Michigan	66.73	33.06	33.27	3.62
Rhode Island	68.16	31.71	31.84	7.21
Wyoming	68.28	31.59	31.72	7.71
Oklahoma	69.45	29.84	30.55	6.14
Alaska	69.49	28.56	30.51	7.02
West Virginia	70.14	28.27	29.86	6.33
Georgia	70.29	26.25	29.71	5.86
Ohio	70.36	29.43	29.64	3.27
Indiana	70.53	29.25	29.47	5.36
Virginia	70.87	28.34	29.13	5.92
District of Columbia	70.96	28.44	29.04	6.69
Delaware	70.97	28.62	29.03	6.80
Arkansas	71.31	24.83	28.69	6.24
North Carolina	71.75	23.76	28.25	3.30
Vermont	71.80	28.34	28.20[c]	7.01
Maryland	72.17	27.76	27.83	5.71
Connecticut	72.30	27.52	27.70	5.83
Florida	72.34	26.69	27.66	3.05
Massachusetts	72.57	27.21	27.43	3.33
Oregon	72.72	27.22	27.28	7.16
Maine	72.89	27.02	27.11	6.21
Mississippi	72.99	23.61	27.01	6.11
Colorado	73.11	26.64	26.89	6.66
Minnesota	73.24	26.68	26.76	6.23
Texas	73.27	25.76	26.73	3.29
Washington	73.51	26.44	26.49	6.54
Tennessee	73.52	24.50	26.48	6.09
New Hampshire	73.81	26.05	26.19	6.99
South Dakota	73.83	26.13	26.17	5.58
Montana	73.91	25.43	26.09	6.22
Alabama	74.19	22.71	25.81	6.87
New York	74.21	25.46	25.79	2.40
Illinois	74.23	25.68	25.77	3.30
Kansas	74.42	25.54	25.58	6.02
Pennsylvania	74.44	25.30	25.56	3.03
Arizona	74.65	24.82	25.35	6.87
Louisiana	74.76	23.59	25.24	6.22
Hawaii	75.39	23.84	24.61	6.32
North Dakota	75.44	24.45	24.56	5.76
South Carolina	75.50	21.14	24.50	5.55
New Jersey	75.67	24.27	24.33	3.07
Wisconsin	75.68	24.40	24.32[c]	6.05
Missouri	75.99	23.92	24.01	5.16
New Mexico	76.21	23.70	23.79	6.10
Iowa	77.23	22.63	22.77	5.92
California	77.71	22.05	22.29	2.43
Nebraska	77.94	21.99	22.06	5.63
Idaho	78.93	20.90	21.07	6.00
Utah	89.15	10.55	10.85	4.66

[a]Crude rates. [b]Confidence intervals are computed for the rate of use of at least one tobacco product. [c]These states were found to have slightly higher rates (less than a difference of 0.25%) of cigarette use than of use of one tobacco product. This finding, which was considered "impossible" by our definition of a person's "using at least one tobacco product," is attributed to the fact that individuals with missing data on one or more tobacco products were excluded from the estimates of using at least one tobacco product.

rette smoking habits between 1955 and 1966. *Vital and Health Statistics*. Series 10, no. 59. Rockville, Md: National Center for Health Statistics; April 1970.

3. Fiore MC, Novotny TE, Pierce JP, Hatziandreau EJ, Patel KM, Davis RM. Trends in cigarette smoking in the United States: the changing influence of gender and race. *JAMA*. 1989;261:49–55.
4. National Center for Health Statistics. *Health, United States, 1987*. Hyattsville, Md: US Dept of Health and Human Services, Public Health Service; September 1989. DHHS publication PHS 89–1597.
5. Marcus AC, Shopland DR, Crane LA, Lynn WR. Prevalence of cigarette smoking in the United States: estimates from the 1985 Current Population Survey. *J Natl Cancer Inst*. 1989;81:409–414.
6. Marcus AC, Crane LA, Shopland DR, Lynn WR. Use of smokeless tobacco in the United States: recent estimates from the Current Population Survey. In: Boyd GM, Darby CA. *Smokeless Tobacco Use in the United States*. Bethesda, Md. 1989:17–23. Natl Cancer Inst Monograph 8.
7. Hanson RH. The Current Population Survey: Design and methodology. Washington, DC: US Dept of Commerce, Bureau of the Census; July 1985. Technical paper 40.
8. Schoenborn CA, Boyd GM. Smoking and other tobacco use: United States, 1987. *Vital and Health Statistics*. Series 10, no. 169. Hyattsville, Md: National Center for Health Statistics; September 1989.
9. Gallup Poll. "Majorities Now Support Ban on Public Smoking, Cigarette Ads." Princeton, NJ: Gallup Organization; September 1988.
10. US Dept of Health and Human Services. *Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General*. Rockville, Md: US Dept of Health and Human Services, Public Health Service; 1989. DHHS publication PHS 89-8411.
11. Pierce JP, Fiore MC, Novotny TE, Hatziandreau EJ, Davis RM. Trends in cigarette smoking in the United States: projections to the year 2000. *JAMA*. 1989:6–71.
12. US Dept of Health and Human Services. *Healthy People 2000: National Health Promotion and Disease Prevention Objectives*. Washington, DC: US Department of Health and Human Services. 1991. DHHS publication PHS 91–50212.

Handmade Cigarettes: It's the Tube That Counts

ABSTRACT

Thirty-one Canadian brands of fine-cut tobaccos for roll-your-own cigarettes (RYOs) were evaluated under standard conditions using mandated tube and filter combinations. Results indicate no evidence of a significant difference in the amounts of tar, nicotine, and carbon monoxide produced by the 31 brands. In addition, the data emphasize that it is primarily the tube and filter combination that determines delivery of toxic constituents. (*Am J Public Health.* 1992;82:107–109)

Murray J. Kaiserman, PhD, and William S. Rickert, PhD

Introduction

There is strong evidence suggesting that many cigarette smokers, being addicted to nicotine, would rather pay a higher price than give up smoking.[1] In countries such as Canada, which has two different taxation rates for manufactured cigarettes and manufactured (fine-cut) tobacco, smokers have the added choice of a cheaper substitute: roll-your-own (RYO) cigarettes.

While most people's image of RYOs is derived from Hollywood westerns, the truth is different. RYOs are easy to make and, after the recent introduction of some new products in Canada, virtually indistinguishable from machine-manufactured cigarettes. RYOs are so popular in Canada that, by the end of 1989, Canadian RYO sales accounted for approximately 14% of the cigarette market.[2]

As smokers of manufactured cigarettes have known for years, cigarettes produce tar and nicotine in measurable and consistent quantities.[1,3,4] Although the quantities measured will vary depending on how the cigarette is smoked,[1,3] smokers are very much aware of these toxic constituents.[1,3,4]

The situation for RYO tobaccos is different. Until recently, there was no information about levels of any toxic constituents on packages of RYO tobaccos. As a result, significant numbers of smokers have not had access to important information about the toxic constituents in the products they consume. Although some information is available, it is primarily in the scientific literature, and thus it is virtually inaccessible to the average smoker.[5]

To rectify this situation, one provision of the Canadian Tobacco Products Control Act, implemented in January 1989, requires manufacturers of cigarette tobaccos for RYOs to measure and report tar, nicotine, and carbon monoxide (CO) yields.[6] The Tobacco Products Control Regulations (TPCR) describe the standard methods to be used.[7]

By October 31, 1989, tar, nicotine, and CO values began appearing on all packages of RYO tobaccos sold in Canada. To determine how the choice of filter and tube combination might affect yields, the Health Protection Branch of Health and Welfare Canada commissioned a study to evaluate differences in the toxic constituent yields of 31 leading Canadian brands of RYO tobacco using the mandated tube and of 17 commercially available, nonmandated filter and tube combinations. All were filled with 900 mg of a homogeneous tobacco blend.

Methods and Apparatus

RYO tobaccos were purchased at either the wholesale or the retail levels in Toronto and Montreal during December 1989. The standard homogeneous tobacco blend was donated by Imperial Tobacco Limited.

Mandated tubes[7] were obtained directly from RJR-MacDonald Inc. of Montreal. Nonmandated tubes (in the 17 commercial brands) were purchased from a wholesaler in Kitchener, Ontario, Canada.

RYOs were prepared, smoked, and analyzed according to the TPCR[7] as amended on July 13, 1989. Smoking parameters included puff volume of 35 ± 0.2 mL; puff frequency of one puff every 60 ± 1 second; puff duration of 2 ± 0.2 second, and butt length of 23 mm.

Results

The mean analytical results (20 observations/brand) for tar, nicotine, and CO for the 31 brands of cigarette tobaccos tested using standard conditions are presented in Table 1. The overall means in milligrams per cigarette are 15.47 (standard deviation [SD] ± 0.41) for tar, 1.14 SD ± 0.09) for nicotine, and 17.69 (SD ± 0.56) for CO. The ranges in milligrams per cigarette are 14.64 to 16.29 for tar, 0.96 to 1.44 for nicotine, and 16.53 to 18.79 for CO. Statistical analyses of these results using Scheffe's

Murray J. Kaiserman is with the Environmental Health Directorate of Health and Welfare Canada in Ottawa, and William S. Rickert is with Labstat Incorporated in Kitchener, Ontario.

Requests for reprints should be sent to Murray J. Kaiserman, PhD, Tobacco Products Section, Health and Welfare Canada, Ottawa, Ontario, Canada K1A OL2.

This paper was submitted to the journal March 14, 1991, and accepted with revisions July 9, 1991.

Editor's Note. See related editorial by Warner and Slade on page 17.

TABLE 1—Mean Analytical Values for Tar, Nicotine, and CO (by Milligrams per Cigarette) for 31 Brands of RYO Tobaccos Tested Using Standard Tubes

Brand Name	Tar	Nicotine	CO
Belvedere	14.79	1.09	17.76
Belvedere Extra Mild	15.49	1.09	16.97
Black Cat	15.37	1.19	18.79
Cameo Menthol	15.74	1.13	18.44
Craven A	15.26	1.14	17.37
Craven A Light	15.32	1.12	17.76
Craven A Special Mild	14.86	1.09	17.03
Craven Menthol	15.6	1.15	17.96
Craven Menthol Special Mild	15.61	1.12	17.63
Daily Mail	15.61	1.31	17.02
Export	15.83	1.17	18.17
Export Extra Light	14.97	1.06	18.03
Export Medium	15.94	1.2	17.57
Export Mild	15.32	1.06	17.26
Export Ultra Light	15.66	1.15	18.03
Mark Ten	15.03	1.05	17.23
Mark Ten Light	15.06	1.03	16.92
Matinee	15.62	1.14	18.47
Matinee Extra Mild	14.64	0.96	17.75
Number 7	16.03	1.16	16.53
Peter Jackson	15.55	1.15	18.25
Players	16.29	1.27	17.93
Players Extra Light	15.0	1.0	17.89
Players Light	15.39	1.14	17.76
Rothmans	15.6	1.13	18.08
Rothmans Extra Light	15.98	1.15	16.96
Rothmans Special Mild	15.6	1.15	17.63
Sportsman Extra Mild	16.1	1.15	17.67
Vogue Mild	14.96	1.13	17.64
Woolco	15.27	1.14	18.62
Zig Zag	16.04	1.44	17.2
Means	15.47	1.14	17.69

TABLE 2—Mean Analytical Values for Tar, Nicotine, and CO (by Milligrams per Cigarette) for 17 Commercially Available Filter Tubes Containing a Standard Blend of Tobacco

Tube Name	Tar	Nicotine	CO
Black Cat RS FT	16.18	1.11	16.52
Century RS FT	18.99	1.25	18.59
Craven A Light KS FT	8.76	0.72	7.33
Embassy KS FT	20.83	1.42	20.28
Export KS FT	13.77	0.98	16.98
Export Light KS FT	14.96	1.03	16.65
Export Ultra Light KS FT	7.93	0.67	8.05
Export Extra Light KS FT	12.97	0.92	14.74
Players KS FT	20.43	1.4	19.79
Players Light KS FT	18.87	1.28	20.1
Players Extra Light KS FT	10.06	0.81	11.08
Rothmans Extra Light KS FT	12.2	0.98	10.11
Rothmans KS FT	20.77	1.4	20.49
Premier KS FT	21.04	1.5	19.1
Woolco RS FT	18.81	1.25	18.0
Woolco KS FT	20.09	1.32	18.18
Woolco Light RS FT	17.31	1.17	16.39

Note. KS = king size; RS = regular size; FT = filter tip.

Multiple Range Test did not provide any evidence for a difference in tar yields.

Table 2 presents analytical results for 17 filter and tube combinations using a standard blend of RYO tobacco. Mean values in milligrams per cigarette range from 7.93 to 21.04 for tar, 0.67 to 1.5 for nicotine, and 7.33 to 20.49 for CO. Within brand families such as Export, Players, and Rothmans, the delivery of tar decreased by about 50% between the "regular" tube and the lowest "ultra light" or "extra light" tube. The decrease in tar delivery between the regular and next lowest light tubes, however, is not as dramatic. In fact, in the case of Export and Export Light, the tar delivery increases although the difference may not be statistically significant.

Discussion

Cigarettes are composed of three major components: tobacco, filters, and the papers used in wrapping the tobacco and filter.[4] Cigarette manufacturers are able to adjust the tar, nicotine, and CO deliveries by manipulating any or all of these components.[3,4] Most often this involves using porous papers and ventilated filters, which dilute the mainstream smoke and thereby reduce the yields of toxic constituents.[3] The TPCR, in mandating a method for testing RYO tobaccos, have ensured that all brands are evaluated on a comparable basis with tobacco as the only variable.

Not surprisingly, there seems to be no substantial differences in RYO tobaccos marketed under the 31 brand names listed in Table 1. Virtually all of the tobacco used by Canadian manufacturers is domestically grown and flue cured, and approximately 89% of the flue-cured tobacco is grown in Ontario.[8] In addition, manufacturers claim that most of their cigarettes are additive free.[9] It is reasonable to conclude, therefore, that there should be no major differences in the standard yields.

Depending on the choice of tube, a smoker can vary tar delivery by approximately 60%. The tubes, however, are sold either separately or in combination with the corresponding tobacco. There is strong evidence that smokers and nonsmokers alike believe that low-yield cigarettes actually do produce lower concentrations of toxic materials.[1,3,10] Although no comparable evidence exists for consumer opinions concerning RYO tobaccos, it is reasonable to assume that users of these products may have a similar impression. Therefore, smokers may be buying low-yield RYO tobaccos in the mistaken belief that it is the tobacco that controls the yields of toxic constituents. But, as the results of this work indicate, it is the tube and filter combination, not the tobacco, that controls the delivery of toxic constituents to smokers. Thus, to achieve the low-yield effect, smokers should be advised to make their choices based on the tube and filter combinations rather than on the tobacco. □

References

1. US Department of Health and Human Services and Centers for Disease Control. *Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General.* Washington, DC: US Dept of Health and Human Services, Public Health Service, and Atlanta, Ga, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; DHHS publication CDC 89-8411.
2. Kaiserman MJ, Allen TA. Trends in Canadian tobacco consumption, 1980–89. *Chronic Diseases in Canada.* 1990;11:54–55.
3. Rickert WS. "Less hazardous" cigarettes: fact or fiction? *NY State J Med.* 1983;83: 1269–1272.
4. R. J. Reynolds Tobacco Company. *Chemical and Biological Studies on New Cigarette Prototypes That Heat Instead of Burn Tobacco.* Winston-Salem, NC: 1988.
5. Rickert WS, Robinson JC, Bray DF, Rogers B, Collishaw NE. Characterization of tobacco products: a comparative study of the tar, nicotine and carbon monoxide yields of cigars, manufactured cigarettes, and cigarettes made from fine-cut tobacco. *Prev Med.* 1985;14:226–233.
6. Tobacco Products Control Act. Can. Rev. Stat. ch. 20 (1988).
7. Tobacco Products Control Regulations. Canada Gazette, part II; 123(2); SOR/DORS/89–21 (December 27, 1988), *amended by* 123(11); SOR/DORS/89–248 (May 4, 1989).
8. Canadian Tobacco Manufacturers' Council. Tobacco in Canada 1989. Ottawa; April 20, 1990.
9. Jalsevac P. Tobacco list names 500 possible additives. *The Gazette* (Montreal). February 11, 1990:A–10.
10. Rickert WS, Robinson JC, Lawless E. Limitations to potential users for data based on the smoking machine of cigarettes. *Nicotine Smoking and the Low Tar Programme.* Independent Scientific Committee on Smoking and Health Symposium. London, England: Oxford Press; 1988:85–89.

Improvement in Intergenerational Health

ABSTRACT

Differences in health status between subjects, their parents, and their children were analyzed in 2206 subjects who had attended the University of Pennsylvania during the 1939 to 1940 school year. Subjects compared their overall health status at the average age of 70 with that of their same-sex parent at the same age and with that of their same-sex child at the approximate average age of 45, providing reasons for reported differences. Thus, health status in family members of the same sex at the same age in 1988 was compared with that in approximately 1963. Subject health was strikingly improved compared with that of their parents a generation earlier, with 58% reporting their health to be better or much better, and only 9% reporting it to be worse or much worse ($P < .001$). The major reasons for the difference were decreased prevalence of chronic conditions and healthier life-styles. The same results were observed in a community-based population of 327 subjects and, even more strikingly, in a group of 422 aging long-distance runners. These observations suggest substantial improvement in senior health status over the past quarter century in selected populations, and they contrast with equivocal changes that have been noted with traditional serial survey techniques. (*Am J Public Health.* 1991;81:109–112)

James F. Fries, MD, Catherine A. Williams, MA, and Dianne Morfeld

Introduction

Has the health of seniors improved over recent years? This controversial question[1–9] is of obvious importance to health policy. Intuitively, it seems that age-specific health status must be improving due to medical advances, decreases in some health risk behaviors, and improvement in mortality rates. But reliable data in this regard are notably lacking, and many studies suggest negative trends.[1] Disability has reportedly increased,[2–5] and prevalence may have increased for heart disease, hypertension, and arthritis.[3,6] Although global health measures remain stable,[1,7,8] death certificate studies show a growing number of listed causes.[9] However, there also has been greater disease ascertainment.[1] Public attitudes may now tolerate lesser degrees of disability, the "paradox of health."[10] There also have been changes in diagnostic standards, hospital admission standards, nursing home beds, and disability reimbursement. Thus, it is not surprising that there is controversy. We present a new analytic approach: we directly compare subject health with parent health.

Methods

We are studying, longitudinally, 2451 subjects who had attended the University of Pennsylvania during the 1939 to 1940 school year to identify risk factors for osteoarthritis.[11,12] Subject age averaged 69.7 years, 70% of subjects were male, and subject education level averaged 17.5 years. It may be of interest that 37% were former smokers but only 2.6% were current smokers.

In 1988 we asked the following questions:

1. How does your *present health* compare with that of your parent of the same sex when he or she was *your present age?* What are the major differences?
2. How does the *present health* of your oldest child of your sex compare with your health when you were *the same age?* What are the major differences?

Thus, we queried subjects for intergenerational health differences between themselves and both their parent at an average age of 70 and their child at age 45, assuming a 25-year generation. We thus compared health status in 1988 with that in

The authors are with the Department of Medicine, Division of Immunology and Rheumatology, Stanford University School of Medicine, Stanford, CA.

Requests for reprints should be sent to James F. Fries, MD, Associate Professor of Medicine, Stanford University School of Medicine, 1000 Welch Road, Suite 203, Palo Alto, CA 94304-1808.

This paper was submitted to the journal June 20, 1990, and accepted with revisions May 6, 1991.

TABLE 1—Characteristics of Subgroups among University of Pennsylvania Students

				Educational Level		Age, y		
	Total	Male	Female	<16	>17	<66	66–70	>70
Total respondents	2451	1891	560	989	1180	255	1537	659
Age, y	69.7	69.6	69.9	68.8	70.2	65.3	68.4	74.5
Male, %	70.0	100.0	0.0	80.0	75.0	76.0	77.0	78.0
Smokers, %	2.6	2.5	3.0	2.9	2.7	3.5	2.5	2.6
Family history of arthritis, %	41.0	39.0	49.0	46.0	44.0	38.0	41.0	42.0
Educational level (years)	17.5	17.5	17.3	15.8	18.8	17.6	17.1	18.2
Disability[a]	0.103	0.079	0.183	0.093	0.104	0.056	0.084	0.167
Global health	82.1	82.2	81.8	82.0	82.1	82.2	82.7	80.6

[a]Based on the Health Assessment Questionnaire.[16–18]

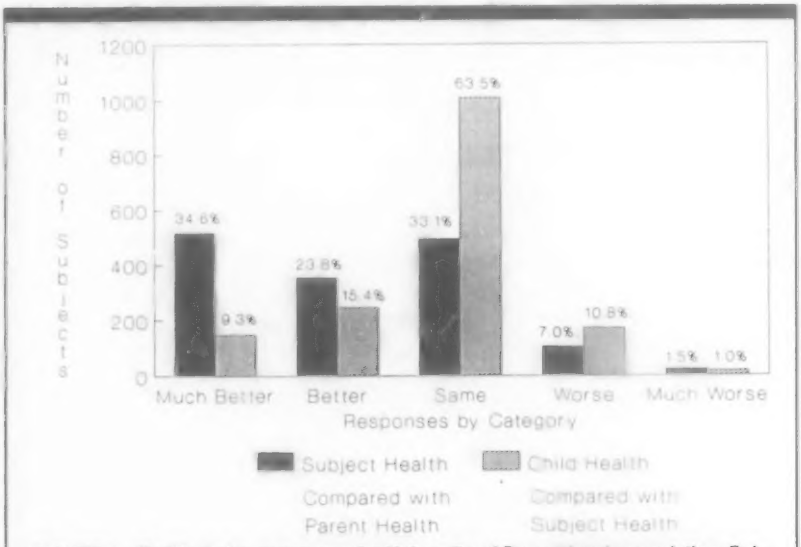


FIGURE 1—Health status responses for University of Pennsylvania population. Subject's average age is 70 years. Numbers represent the percentage of subjects in each category.

approximately 1963. This intrafamily comparison controls for age, sex, and, in part, genetic and social factors.

We also studied two additional groups at similar ages: 422 long-time, long-distance runners from a 50-plus runners club and 317 community controls from the Lipid Research Clinic Study.[13–15]

Independent variables studied included sex, highest grade level, age, race, smoking history, family history of arthritis, disability index (0 to 3) using the Health Assessment Questionnaire (HAQ),[16–18] and global health as marked on a horizontal analog scale (0 to 100)[16,18] Dependent variables were subject health compared with that of both parent and child, analyzed by chi-square analysis. Subject characteristics are summarized in Table 1.

Results

Main results are shown in Figure 1. An overwhelming 58% of subjects reported that their health was better or much better than that of their parent compared with 9% reporting health that was worse or much worse ($P < .0001$). The mode was "much better." In contrast, when comparing their health with that of their same-sex child, 25% reported the child's health to be better or much better and 12% reported it to be worse or much worse ($P < .001$). Median and mode were "same."

Table 2 summarizes responses in subgroups characterized by sex, age, and educational level, with categories of better/much better and categories of worse/much worse collapsed. The same results obtained in every subgroup (all $P < .001$). Among the 50-plus runners club subjects, 82% reported better health than their parent ($P < .0001$), as did 66% among the community control population ($P < .001$). In neither of these populations were comparisons with child health statistically different.

Reported reasons for better health are grouped in Table 3. Presence of cardiac, respiratory, and other diseases in the parent accounted for approximately half of the responses. Greater exercise, life-style and diet factors, obesity, and smoking and drinking accounted for the remainder.

Discussion

Subjects reported marked health improvement compared with their parents. Results were confirmed in multiple subgroups and in two additional populations. The attributed reasons for improvement are consistent with decreases in cardiovascular disease and improvements in lifestyle between 1963 and 1988.

Are there intrinsic biases in this technique? Three general caveats apply. First, in cross-sectional studies there is selection bias (left-censoring) due to deaths prior to study date in the original cohort. Subjects thus may have better health than average. However, we did not analyze subjects outliving parents and thus in "much better" health; if we had, results would have been even more pronounced. Additionally, because of increased longevity, fewer subjects than parents should be left-censored by death. This potential bias, therefore, works against the observed results.

Second, results are based on self-report. There could be a general tendency to see personal health as better than it is. This potential bias cannot be excluded, but the neutral or negative responses to the "control" comparison with child health tend to make this less likely. The "paradox of health" suggests that individ-

TABLE 2—Intergenerational Health among University of Pennsylvania Students

				Educational Level		Age, y		
	Total	Male	Female	<16	>17	60–65	66–70	>70
Total respondents	2206	1699	507	973	1159	230	1401	575
Parent died	711	587	124	312	370	39	431	241
Subject vs parent	1495	1112	383	661	789	191	970	334
% Better	58	56	65	56	60	60	59	57
% Same	33	36	26	36	31	30	33	34
% Worse	9	8	9	8	9	10	8	9
Total respondents	2112	1650	462	942	1101	223	1359	530
Child died	37	32	5	18	19	3	26	8
No child	487	339	148	199	266	44	289	154
Subject vs child	1588	1279	309	725	816	176	1044	368
% Better	25	25	23	24	26	20	25	25
% Same	63	65	58	64	63	64	63	65
% Worse	12	10	19	12	11	16	12	10

TABLE 3—Reported Reasons for Better Health among University of Pennsylvania Students (by percent)

				Educational Level		Age, y		
	Total	Male	Female	<16	>17	57–65	66–70	>70
Total respondents	(407)	(269)	(138)	(172)	(218)	(64)	(224)	(119)
Health habits								
Exercise	25.3	25.3	25.4	27.3	23.4	28.1	23.7	26.9
Life-style and diet	5.2	5.2	5.1	3.5	6.4	4.7	4.0	7.6
Weight	5.4	5.9	4.3	5.2	6.0	4.7	6.3	4.2
Smoke/drink	2.5	3.7	0.0	4.7	0.9	3.1	2.7	1.7
Diseases								
Cardiac	22.6	24.2	19.6	23.8	19.7	15.6	27.2	17.6
Respiratory	2.7	3.0	2.2	1.7	3.7	6.3	1.8	2.5
Other major disease[a]	28.0	26.4	31.2	21.5	33.9	25.0	25.4	34.5
Other minor disease[b]	8.4	6.3	12.3	12.2	6.0	12.5	8.9	5.0

[a]E.g., strokes, cancers, neurological disorders, arthritis, diabetes.
[b]E.g., ulcers, fractures, headaches, hernias.

ual responses might actually be more often pessimistic than optimistic.[10] Previous validation studies on these populations comparing reported to observed disability levels showed equal accuracy in exercising and sedentary groups.[3]

Third, subjects studied were educated and, generally, not underprivileged. Nevertheless, results were consistent across all groups analyzed, including a community-based group. Readers may confirm results simply by posing the same questions to themselves, their family, and their acquaintances.

The compression of morbidity (COM) hypothesis[19–21] holds that, if the average onset age of infirmity is delayed and if this delay is greater than increases in longevity, the amount of life spent infirm will decrease. Such a more "rectangular" course anticipates prolongation of the period of health, followed by a more precipitate decline in health prior to death. These circumstances indirectly support the COM hypothesis, which requires a decrease in age-specific morbidity rates. Adult life expectancy (from age 50) increased by about 3 years between 1963 and 1988.[22] The striking health improvement reported will probably have a greater effect on than this modest longevity increase, but direct comparison of morbidity and mortality trends is not possible from our data.

Nor can the cause of the improvement be established definitively. Life-style changes that were cited suggest a partial role of personal health risk reduction. The decreased frequency of chronic diseases that was cited may be due to prevention or to improved care. These results cannot be generalized to underprivileged populations that experience fewer changes in health risks or have inadequate access to care. However, results suggest that, at least in certain population segments, substantial improvements in senior health status have occurred over the past generation. □

Acknowledgments

This study was supported in part by a grant from the National Institutes of Health (NIH AM-21395) to ARAMIS (Arthritis, Rheumatism, and Aging Medical Information System).

References

1. Verbrugge LM. Recent, present, and future health of American adults. *Ann Rev Public Health.* 1989;10:333–361.
2. Colvez A, Blanchet M. Disability trends in the United States population 1966–76: analysis of reported causes. *Am J Public Health.* 1981;71:464–471.
3. Verbrugge LM. Longer life but worsening health? Trends in health and mortality of middle-aged and older persons. *Milbank Mem Fund Q.* 1984;62:475–519.
4. Chirikos TN. Accounting for the historical rise in work-disability prevalence. *Milbank Mem Fund Q.* 1986;64:271–301.
5. Feldman JJ. Work ability of the aged under conditions of improving mortality. *Milbank Mem Fund Q.* 1983;61:430–444.
6. Feinleib M, Wilson RW. Trends in health in the United States. *Environ Health Perspect.* 1985;62:267–276.

7. Palmore EB. Trends in the health of the aged. *Gerontologist.* 1986;26:298–302.
8. Ycas MA. Are the eighties different? Continuity and change in the health of older persons. In: *Proceedings of the 1987 Public Health Conference on Records and Statistics.* Hyattsville, Md: National Center for Health Statistics; 1988: 57–61. US Dept of Health and Human Services publication (PHS) 88–1214.
9. Israel RA, Rosenberg HM, Curtin LR. Analytical potential for multiple cause-of-death data. *Am J Epidemiol.* 1986;124:161–179.
10. Barsky AJ. The paradox of health. *N Engl J Med.* 1988;318:414–418.
11. Paffenbarger RS, Wing AL, Hyde RT. Chronic disease in former college students, XVI: physical activity as an index of heart attack risk in college alumni. *Am J Epidemiol.* 1978;108:161–175.
12. Whittemore AS, Paffenbarger RS, Anderson K, Lee J. Chronic disease in former college students, XXVII: early precursors of site specific cancers in college men and women. *JNCI.* 1985;74:43–51.
13. Lane NE, Bloch DA, Jones HH, Marshall WH, Wood PD, Fries JF. Long-distance running, bone density, and osteoarthritis. *JAMA.* 1986;255(9):1147–1151.
14. Lane NE, Bloch DA, Wood PD, Fries JF. Aging, long-distance running, and the development of musculoskeletal disability. *Am J Med.* 1987;82:772–780.
15. Michel BA, Bloch DA, Fries JF. Weight-bearing exercise, overexercise, and lumbar bone density over age 50 years. *Arch Intern Med.* 1989;149:2325–2329.
16. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum.* 1980;23:137–145.
17. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. *J Rheumatol.* 1982;9:789–793.
18. Wolfe F, Kleinheksel SM, Cathey MA, Hawley DJ, Spitz PW, Fries JF. The clinical value of the Stanford Health Assessment Questionnaire functional disability index in patients with rheumatoid arthritis. *J Rheumatol.* 1988;15:1480–1488.
19. Fries JF. Aging, natural death, and the compression of morbidity. *N Engl J Med.* 1980;303:130–135.
20. Fries JF. Aging, illness, and health policy: implications of the compression of morbidity. *Perspect Biol Med.* 1988;31(3):407–428.
21. Fries JF. Compression of morbidity: near or far? *Milbank Mem Fund Q.* 1990;67(2): 208–232.
22. US Dept of Health and Human Services. *Life Tables for the United States: 1900–2050.* Washington, DC 1982. Actuarial Study Social Security Administration publication 11–11534.

The Effect of Legal Drinking Age on Fatal Injuries of Adolescents and Young Adults

ABSTRACT

This study examined the effect of legal drinking age (LDA) on fatal injuries in persons aged 15 to 24 years in the United States between 1979 and 1984. Effects on pre-LDA teens, adolescents targeted by LDA, initiation at LDA, and post-LDA drinking experience were assessed. A higher LDA was also associated with reduced death rates for motor vehicle drivers, pedestrians, unintentional injuries excluding motor vehicle injuries, and suicide. An initiation effect on homicides was identified. Reductions in injury deaths related to drinking experience were not found. In general, a higher LDA reduced deaths among adolescents and young adults for various categories of violent death. (*Am J Public Health.* 1992;82:112–115)

Nancy E. Jones, DrPH, CPNP, Carl F. Pieper, DrPH, and Leon S. Robertson, PhD

Introduction

While the effect of raising the legal drinking age (LDA) on reducing the involvement of young drivers in fatal motor vehicle crashes has been demonstrated,[1,2] the potential impact of this legislation on other alcohol-related violent deaths[3–16] has not been examined except in one state.[17] Therefore, this study investigates the relationship between LDA and all fatal injuries among adolescents and young adults.

The literature has suggested four possible effects of LDA:

1. *Legal Effect:* A high LDA reduces fatal injuries among adolescents targeted by the law.[2]
2. *Spillover Effect:* A low LDA increases fatal injuries among adolescents younger than the LDA through earlier access to alcohol from their older aged peers.[1,2]
3. *Initiation Effect:* An increase in fatal injuries in the year that access to alcohol becomes legal offsets the reductions of earlier years.[18,19]
4. *Experience Effect:* Reductions in fatal injuries occur as drinking experience increases. Fatal injuries decrease in the years following legal access.[18,19]

Methods

All 50 states and the District of Columbia were studied for the years 1979 through 1984. In each state and each year, LDA and injury death rates for persons 15 through 24 years of age were determined. Information on LDA was obtained from the Insurance Institute for Highway Safety (Table 1). Mortality data were col-

Nancy E. Jones is with Hunter College, New York. Carl F. Pieper is with the Center for Aging and Human Development, Duke University, Durham, NC. Leon S. Robertson is with Nanlee Research, Branford, CT.

Requests for reprints should be sent to Nancy E. Jones, DrPH, CPNP, Hunter College, Hunter-Bellevue School of Nursing, 425 East 25th Street, New York, NY 10010.

This paper was submitted to the journal September 10, 1990, and accepted with revisions May 2, 1991.

lected from the National Center for Health Statistics. Injury deaths (E800–999) were identified, and six categories of death defined on the basis of the *International Classification of Diseases* (9th version; ICD-9) E-code were examined: motor vehicle driver, motorcyclist, pedestrian, unintentional injury excluding motor vehicle, suicide, and homicide (Table 2). The 1980 census provided population data to compute death rates. Census figures were adjusted by extrapolation to arrive at denominator data for the years 1979 and 1981 through 1984 (e.g., persons who were 18 in 1980 were considered 19 in 1981).

For each of the six categories of injury, the data were aggregated into death rates per 100 000 population by calendar year, age, and LDA, thus forming a matrix of 240 rates (6 calendar years × 10 single years of age × 4 LDAs). Calendar year was added to control for temporal trends in injuries. Logistic regression was used for analysis. The following model was fitted to the data:

$$\text{Rate} = a + b_1(\text{CALYR}) + b_2(\text{AGE}) + b_3(\textit{SPILLOVER}) + b_4(\textit{LEGAL}) + b_5(\textit{INITIATION}) + b_6(\textit{EXPERIENCE})$$

where Rate = fatal injury rate, CALYR = calendar year, AGE = age of the fatally injured person, SPILLOVER = LDA − age, if age < LDA (else 0), LEGAL = 1 if drinking legal at age of death or else 0, INITIATION = 1 if LDA = age or else 0, and EXPERIENCE = age − LDA + 1, if age > LDA (else 0). For the variables calendar year and age, a system of indicator variables was constructed to allow for expected nonlinearities in the death rate over time and age.

The following hypotheses were tested. If a low legal drinking age increases injuries in adolescents younger than the legal minimum—the "spillover effect"—injury rates will decrease as the years before legal access increase. If a high drinking age reduces injuries only among adolescents targeted by the legislation—the "legal effect"—a low drinking age will be associated with a high injury death rate among those who can drink legally. If an "initiation effect" occurs, fatal injuries will increase more than expected in the year that drinking becomes legal. If injuries decrease in the years following legalization—the "experience effect"—injury rates will decrease as experience with alcohol increases beyond the year that drinking becomes legal.

TABLE 1—Legal Drinking Age (LDA) Changes, United States, 1979–1984[a]

State	LDA	Date of Change	State	LDA	Date of Change
Ala	19		Mont	19	
Alaska[b]	19		Neb	19–20	7/19/80
Ariz[c]	19		Nev	21	
Ark	21		NH[d]	20	
Calif	21		NJ	18–19	7/02/80
Colo[d]	21			19–21	1/01/83
Conn	18–19	7/01/82	NM	21	
	19–20	10/01/83	NY	18–19	12/04/82
Del[e]	20–21	1/01/84	NC[e]	18–19	10/01/83
DC	18		ND	21	
Fla	18–19	10/01/80	Ohio[d]	21	
Ga	18–19	9/01/80	Okla[d]	21	
Hawaii	18		Ore	21	
Idaho	19		Pa	21	
Ill[f]	19–21	1/01/80	RI	18–19	7/01/80
Ind	21			19–20	7/01/81
Iowa	19		SC[i]	18–19	1/01/84
Kan[d]	21		SD[d]	21	
Ky	21		Tenn[m]	19	
La	18		Tex	18–19	9/01/81
Me	20		Utah	21	
Md[g]	18–21	7/01/82	Vt	18	
Mass[h]	20		Va[n]	18–19	7/01/83
Mich	21		Wash	21	
Minn	19		WVa	18–19	7/01/83
Miss	21		Wis	18	
Mo	21		Wyo	19	

[a]Laws apply to all alcoholic beverages except where noted.
[b]10/26/83 from 19 to 21; grandfathered persons born on or before 12/31/64.
[c]12/31/84 from 19 to 21; grandfathered persons 19 before 12/31/84.
[d]18 (3.2% beer).
[e]Grandfathered persons 20 on 1/1/84.
[f]From 19 for beer and wine to 21 for all alcohol.
[g]From 18 for beer and light wine to 21 for all alcohol.
[h]4/16/79 from 18 to 20.
[i]5/24/79 from 18 to 20.
[j]1/2/80 from 18 to 19; grandfathered persons 18 before 1/2/80. 1/1/83 from 19 to 21; grandfathered persons 19 or 20 by 12/31/82.
[k]10/1/83 beer and unfortified wine law raised from 18 to 19.
[l]1/1/84 beer and wine law from 18 to 19.
[m]6/1/84 from 19 to 21; grandfathered persons born before 8/1/65.
[n]7/1/83 from 18 for "on-sale" and 19 for "off-sale" to 19 for all beer.

TABLE 2—Injury Death Rates and Deaths by Cause, for Persons Aged 15–24 Years, United States, 1979–1984

Cause of Death	ICD-9E Code	Death[a] Rate	No. of Deaths
Motor vehicle driver	E810–E819 (.0) E822–E824 (.0)	13.48	33 650
Motorcyclist	E810–E819 (.2, .3) E822–E824 (.2, .3)	4.38	10 923
Pedestrian	E810–E819 (.7) E822–E824 (.7)	3.57	8903
Unintentional[b] injuries excluding motor vehicle	E800–E807 E820–E821 E825–E949	15.24	38 034
Suicide	E950–E959	12.24	30 542
Homicide	E960–E978	13.84	34 526
All injuries[c]	E800–E999	82.55	206 002

Note. ICD-9 = *International Classification of Diseases*, 9th revision. (.0) = motor vehicle driver; (.1) = motor vehicle passenger; (.2) = motorcycle driver; (.3) = motorcycle passenger; (.7) = pedestrian; and (.9) = unknown.
[a]Number of deaths per 100 000 persons.
[b]Includes causes of injury death (e.g., drowning, house fire, fall).
[c]Includes the above categories of injury death as well as Motor Vehicle Occupants, E810–E819 (.1, .9) and E822–E824, (.1, .9) and Undetermined, E980–989.

TABLE 3—Estimated Effects of Legal Drinking Age on Injury Death Rates among Younger Adolescents (Spillover), Targeted Adolescents (Legal), Year of Majority (Initiation), and Young Adults (Experience), Aged 15–24, 1979–1984[a]

	Spillover		Legal		Initiaiton		Experience	
Cause of Death	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Motor vehicle driver	0.958[b]	(0.941,0.975)	1.062	(0.999,1.128)	1.009	(0.962,1.059)	1.051[b]	(1.037,1.065)
Motorcyclist	1.030	(0.997,1.063)	1.026	(0.921,1.141	1.080	(0.992,1.177)	0.989	(0.965,1.012)
Pedestrian	0.895[b]	(0.868,0.923)	0.933	(0.828,1.051)	1.013	(0.918,1.117)	1.124[b]	(1.094,1.154)
Unintentional injuries excluding motor vehicles	0.961[b]	(0.947,0.976)	0.982	(0.927,1.040)	0.947[b]	(0.903,0.994)	1.025[b]	(1.012,1.038)
Suicide	1.032[b]	(1.012,1.053)	1.097[b]	(1.028,1.17	1.004	(0.953,1.057)	0.985[b]	(0.972,0.998)
Homicide	1.010	(0.991,1.029)	0.994	(0.935,1.056)	1.051[b]	(1.001,1.104)	1.070[b]	(1.057,1.083)
All injuries	0.993[b]	(0.986,0.999)	1.039[b]	(1.013,1.064)	1.010	(0.990,1.031)	1.021[b]	(1.016,1.027)

Note. OR = odds ratio; CI = confidence interval.
[a]Controlling for age and calendar year.
[b]$P < .05$.

Results

In Table 3 (first column), the coefficients are in the predicted direction (less than one) and significant for the following three categories: motor vehicle driver (odds ratio [OR] = 0.958), pedestrian (OR = 0.895), and unintentional injuries excluding motor vehicle (OR = 0.961). Yearly reductions in deaths of 4.2%, 10.5%, and 3.9%, respectively, are seen. While the coefficients are not in the predicted direction for three categories of fatal injury, only one, suicide, has a 95% confidence interval greater than one. For all injuries combined (OR = 0.993), a reduction in deaths of 0.7% is observed for each year that the legal drinking age is delayed.

In the second column, the coefficients are in the predicted direction (greater than one) for three categories of fatal injury: motor vehicle driver, motorcyclist, and suicide. However, only suicide (OR = 1.097) is significant. This coefficient suggests that among persons of a given age who can drink legally, the suicide rate is 9.7% greater than among persons of the same age who cannot drink legally, controlling for the other effects. For all injuries combined (OR = 1.039), a 3.9% increase in death rate is observed among persons of a given age who can drink legally.

In the third column, while the coefficients are in the predicted direction (greater than one) for five categories of fatal injury, only one, homicide (OR = 1.051), is significant. This coefficient suggests a 5% increase in homicides in the year that access to alcohol becomes legal. For unintentional injuries excluding motor vehicles (OR = 0.947), the coefficient indicates 5% fewer injuries than expected in the first year of legal drinking.

In the fourth column, the coefficients are opposite to the predicted direction (less than one) and significant for motor vehicle driver (OR = 1.051), pedestrian (OR = 1.124), unintentional injuries excluding motor vehicle (OR = 1.025), and homicide (OR = 1.070). The coefficient is in the predicted direction for suicide (OR = 0.985). For all injuries combined (OR = 1.021), a death-rate increase of 2.1% per year post-LDA was found.

Discussion

While the effect of LDA on fatal injuries varies among categories and age groups, the results of this study suggest that the net benefit of a drinking age of 21 is found not only among motor vehicle drivers, but among other categories of violent death as well. The findings indicate that raising the LDA may have three effects: delaying legal access to alcohol among pre-LDA adolescents, preventing traumatic deaths that occur with legal access, and delaying the onset of heavy drinking and associated fatal injuries that can occur with experience.

The results of this study are consistent with the findings of earlier research on LDA and driver fatalities in which alcohol was actually measured.[20] Furthermore, the LDA during the period studied was not related to the percentage of the population residing in metropolitan areas among the states, a potential confounding factor. However, the results would be more definitive if blood alcohol concentrations had been available for all categories of traumatic death. Universal measurement and reporting of blood alcohol in the fatally injured by age would enhance the evaluation of this and other alcohol countermeasures. □

Acknowledgments

This study was supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (1 R01 AA07246–91).

An earlier version of this paper was presented at the annual meeting of the American Public Health Association, Boston, 1988.

References

1. Wagenaar AC. *Alcohol, Young Drivers and Traffic Accidents.* Lexington, Mass: Lexington Books; 1983.
2. Williams A, Zador P, Harris S, Karpf R. The effect of raising the legal minimum drinking age on involvement in fatal crashes. *J Leg Stud.* 1983;XII:169–179.
3. Haddon W, Valien P, McCarroll JR, Umberger CJ. A controlled investigation of the characteristics of adult pedestrians fatally injured by motor vehicles in Manhattan. *J Chronic Dis.* 1961;14:655–678.
4. Baker SP, Robertson LS, O'Neill B. Fatal pedestrian collisions. *Am J Public Health.* 1974;64:319–325.
5. Waller JA. Non-highway injury fatalities: the roles of alcohol and problem drinking, drugs and medical impairment. *J Chronic Dis.* 1972;21:33–45.
6. Waller JA. Alcohol and unintentional injury. In: Kissin B, Begleiter H, eds. *The Biology of Alcoholism.* New York, NY: Plenum Publishing; 1976.
7. Waller JA. Unintentional injury among the medically impaired and elderly. *Public Health Rep.* 1985;100:577–579.
8. Haddon W, Baker SP. Injury control. In: Clark D, MacMahon B, eds. *Preventive and Community Medicine.* Boston, Mass: Little, Brown; 1981.
9. Robertson LS. *Injuries: Causes, Control Strategies, and Public Policy.* Lexington, Mass: Lexington Books; 1983.
10. Baker SP, O'Neill B, Karpf R. *The Injury Fact Book.* Lexington, Mass: Lexington Books; 1984.
11. Baker SP. Injury facts, risk groups, and injury determinants. *Public Health Rep.* 1985;100:581–582.

12. Smith GS. Measuring the gap for unintentional injuries. *Public Health Rep.* 1985;100:565–568.
13. Trumble JG. The 1990 injury prevention objectives—NIAAA's perspective and current status. *Public Health Rep.* 1985;100:618–620.
14. National Research Council and Institute of Medicine, Committee on Trauma Research, Commission on Life Sciences. *Injury in America: A Continuing Public Health Problem.* Washington, DC: National Academy Press; 1985.
15. Smith GS, Kraus JF. Alcohol and residential, recreational, and occupational injuries. *Ann Rev Public Health.* 1988;9:99–121.
16. Council on Injury Control. Injury fact sheets—alcohol and injuries. *Council on Injury Control Quarterly Newsletter.* 1984;1:2.
17. Hingson R, Merrigan D, Heeren T. Effects of Massachusetts raising its legal drinking age from 18 to 20 on deaths from teenage homicide, suicide and nontraffic accidents. *Pediatr Clin North Am.* 1985;32:221–232.
18. Asch P, Levy DT. Does the minimum drinking age affect traffic fatalities? *Journal of Policy Analysis and Management.* 1987;6:180–192.
19. Asch P, Levy DT. *Young Driver Fatalities: The Roles of Drinking Age and Drinking Experience.* New Brunswick, NJ: Rutgers Bureau of Economic Research; 1987.
20. Robertson LS. Blood alcohol in fatally injured drivers and the minimum legal drinking age. *Journal of Health Politics, Policy and Law.* 1989;14:817–825.

Evaluating the Treatment of Sexually Transmitted Diseases at an Urban Public Hospital Outpatient Clinic

Paul G. Shekelle, MD, MPH, and Jacqueline Kosecoff, PhD

ABSTRACT

Sexually transmitted diseases occur disproportionately among the poor, are often treated in public hospitals and clinics, and have not been subjected to quality-of-care evaluation. We designed a medical record abstraction system using well-established, specific process-of-care criteria drawn from the medical literature and experts and grouped into three levels of quality: excellent, adequate, and minimal. One hundred seventy-six consecutive patients were identified from the clinic logbook and their medical records abstracted. Deficiencies in history taking, physical examination, laboratory testing, treatment, and public health reporting were identified. (*Am J Public Health.* 1992;82:115–117)

Introduction

Gonorrheal urethritis, nongonorrheal urethritis, and pelvic inflammatory disease are among the most commonly occurring sexually transmitted diseases. Appropriate treatment is curative, while lack of appropriate treatment can lead to serious implications. Sexually transmitted diseases occur disproportionately among the poor and underinsured,[1] who are more likely to receive care at a public hospital or a health department clinic. It is important to understand the quality of care delivered at these institutions. Therefore, we undertook an evaluation of the treatment of sexually transmitted diseases at one urban public hospital clinic.

Methods

The study hospital is a county-run facility in the Los Angeles metropolitan area that serves a poor, predominantly Hispanic population. The hospital's medical walk-in clinic sees adult patients with nonemergent problems and is staffed by full-time and part-time physicians and by medical residents. The nursing staff is bilingual. Physicians do not use disease-specific forms to record clinical information. Virtually all patients with sexually transmitted diseases presenting Monday through Friday are seen in this clinic.

We designed a medical record abstraction system to measure the quality of the process of care given to patients diagnosed with gonorrhea, nongonorrheal urethritis, and pelvic inflammatory disease. In order to develop explicit process-of-care criteria, we examined recommendations from the Centers for Disease Control, the Los Angeles County Public Health Department, and the scientific literature, and we also talked with experts. The final set of criteria was approved by the clinic director and the physician in charge of quality assurance for the treatment of sexually transmitted diseases. The criteria were grouped into three levels of quality: excellent, adequate, and minimal (Table 1).

Patients presenting to the clinic from November 1987 through February 1988 were eligible for the study. Patients who had a physician-recorded diagnosis that was compatible with one of the study diseases were included. One hundred seventy-six patients for whom the medical record was available were identified, and their medical records were abstracted. The physician-recorded history and physical, progress notes, physician orders, nursing notes, and laboratory reports were reviewed. Analytic methods included frequency tables and *t* tests.

At the time of the study the authors were with the UCLA Division of General Internal Medicine, Department of Medicine, and the UCLA Department of Health Services, School of Public Health. Dr. Kosecoff is also with Value Health Sciences, Santa Monica, Calif.

Requests for reprints should be sent to Paul G. Shekelle, MD, MPH, Department of Medicine, VAMC West Los Angeles, Wilshire and Sawtelle Blvds., Los Angeles, CA 90073.

This paper was submitted to the journal January 14, 1991, and accepted with revisions May 28, 1991.

TABLE 1—Explicit Criteria for Treatment of Sexually Transmitted Diseases

Criteria	Quality[1] E	A	M	Records Applicable No. of Charts	% Compliance
History					
Documentation of:					
Presence or absence of:					
Symptoms	x	x	x	176	100
Nongenital exposure	x			176	2
Known STD exposure	x			176	50
Last menstrual period	x	x	x	54	59
Last sexual contact	x			176	35
Number of sexual contacts in last 30 days	x			176	27
Any mention of sexual partner(s)		x		176	58
Physical examination					
Documentation of:					
Vital signs	x	x	x	176	99
Men: Exam of penis	x	x	x	122	98
testes	x	x		122	31
inguinal nodes	x	x		122	42
Women: Pelvic exam	x	x	x	54	94
Nongenital exam (if exposed)	x			3	67
Laboratory examination					
Documentation of:					
Gram stain if indicated[a]	x			74	49
Gonorrheal culture	x	x	x	176	87
Chlamydia test if indicated[b]	x			158	50
Blood syphilis serology	x	x		176	71
Pregnancy test if indicated[c]	x	x	x	16	44
Diagnosis	x	x	x	176	100
Approved treatment given[d]	x	x	x	176	88
Follow-up					
Public health report filed	x	x		79	27
Follow-up referral given	x	x	x	176	95
Follow-up appointment made				168	28
Follow-up appointment kept				47	89

Note. E = excellent; A = adequate; M = minimal; x denotes criteria needed to satisfy that level of quality.
[a]Gram stain indicated in male patients with discharge.
[b]Chlamydia test indicated in all patients with discharge.
[c]Pregnancy test indicated if there was uncertainty regarding menstrual history.
[d]Approved treatment conformed to Centers for Disease Control guidelines.

TABLE 2—Process-of-Care Deficiencies Resulting in a Classification of Less-than-Minimal Quality for 49 of 176 patients

Process Criteria	No. of Records Deficient[a]
Vital signs	1
Physical examination	5
Gonorrheal culture ordered	23
Pregnancy test (when indicated)	9
Approved treatment	22
Follow-up referral made	8

[a]Some records failed for more than one reason.

Results

Patient Characteristics

Sixty-nine percent of the 176 patients were males and 78% were age 30 or younger. One hundred fifty-eight (90%) of the patients presented because they were symptomatic; the remainder were asymptomatic patients with a known exposure. Of medical records that contained information about sexual contacts in the previous 30 days, 9 records (19%) indicated that the patient denied any recent sexual encounter, 33 records (69%) indicated that the patient had one partner, 4 records (8%) indicated two partners, and 2 records (4%) indicated three partners.

Compliance with Quality-of-Care Criteria

The frequency with which specific process-of-care items were recorded is shown in Table 1. For example, 100% of the records had documentation of the presence or absence of symptoms, 2% had documentation of the physician having asked about nongenital exposure site, 35% had documentation of the last sexual contact, and 58% had documentation of the patient's last sexual partner.

Two records (1%) met the criteria for excellent quality, 19 records (11%) met the criteria for adequate quality (but not excellent quality), and 106 records (60%) met the criteria for minimal quality (but not excellent or adequate quality). In total, 49 records (28%) failed to meet the minimal-quality criteria. The reasons for failure to meet the minimal level of quality are shown in Table 2. In addition, patients with nongonorrheal urethritis were more likely to receive less-than-minimal-quality care than were those with gonorrhea (41% and 19%; $P < .05$).

Discussion

Our study found clinically meaningful deficiencies in the treatment of sexually transmitted diseases at one urban public hospital. These deficiencies included underuse of gram stains, chlamydia assays, and blood syphilis serology testing; inadequate history taking, patient risk assessment, physical examination, and patient follow-up; improper treatment; and poor public health reporting.

Some of these deficiencies may reflect local problems, for instance, a 2-hour turnaround time to receive results of a gram stain and a policy that a patient should call to schedule a return appointment rather than be given a follow-up appointment before leaving the clinic. Chlamydia assays, on the other hand, may have been underused because physicians doubted the worth of this test. The low rate of positive results (13% of chlamydia assays demonstrated positive results in this study, which is less than half the 31% rate of positive gonorrheal cultures) was known to the clinic staff. In addition, national recommendations for proper ther-

apy for all patients with gonorrhea include doxycycline for possible coexistent chlamydia. Removing the chlamydia test criterion from the definition of excellent quality increases membership in this category from two to six patients. Finally, the low level of public health reporting and the underuse of syphilis serology testing is particularly important since contact tracing is crucial to prevent further spread of disease.[2] In the study population, 3% of the syphilis serology tests were positive.

A goal of public health policy is to encourage poor patients with sexually transmitted diseases to seek care. We have shown that when such patients did seek care at one outpatient clinic, significant deficiencies existed in the documentation of well-established process-of-care criteria. If our results are generalizable, then the goals of public health policy should be broadened to include improving the quality of care provided to patients with sexually transmitted diseases. □

Acknowledgments

This paper was presented in part at the joint session of the national meetings of the Society of General Internal Medicine and the American Federation for Clinical Research in Seattle, May 1991.

The authors thank Dr. Satwant Sidhu, MD, MSPH, and Dr. Pramod Gan, MD, for their assistance in defining the process-of-care criteria used in this study and Dr. Robert Brook, MD, ScD, for his valuable critique of the manuscript.

References

1. Aral SO, Holmes KK. Epidemiology of sexual behavior and sexually transmitted diseases. In: Holmes KK, Mardh P, Sparling PF, et al., eds. *Sexually Transmitted Diseases*. New York, NY: McGraw-Hill; 1990:19–22.
2. Cates W, Meheus A. Strategies for development of sexually transmitted diseases control programs. In: Holmes KK, Mardh P, Sparling PF, et al., eds. *Sexually Transmitted Diseases*. New York, NY: McGraw-Hill; 1990:1023–1029.

Surveillance of Sentinel Occupational Mortality in the District of Columbia: 1980 to 1987

Ann Cottrell, BA, Eugene Schwartz, MD, MPH, Rosemary Sokas, MD, MOH, Vincent Kofie, PhD, and Laura Welch, MD

ABSTRACT

Epidemiological surveillance of sentinel occupationally related deaths commonly relies on computerized analyses of mortality data obtained from vital statistics records. A computer search of death records in the District of Columbia for the period 1980 to 1987 identified 15 cases that noted asbestosis, silicosis, coal worker's pneumoconiosis, or primary cancer of the pleura/mesothelioma as the underlying cause of death. A manual review of the death certificates for the same period identified three times as many cases (n = 48) with any mention of these conditions. Problems with performing surveillance of these events using death certificates include the lack of sufficient information to identify mesotheliomas and the failure to code and computerize all contributing causes of death. (*Am J Public Health*. 1991;82:117–119)

Introduction

Occupational disease surveillance can be used to plan, implement, and evaluate public health interventions that are designed to prevent or control work-related morbidity and mortality. The simplest surveillance systems have been based on the concept of "sentinel health events," as described by Rutstein et al. and others.[1,2]

An ongoing surveillance system may serve to (1) identify occupational or community cohorts that may benefit from medical screening, counseling, or epidemiological follow-up; (2) identify industries, particular companies, or neighborhoods that may require industrial hygiene evaluation, engineering controls, or hazardous waste remediation; (3) measure the impact of selected work-related exposures on the mortality of a population; and (4) provide an evaluation tool to determine the impact of regulatory measures.[1] As of 1988, 40 states and the District of Columbia have initiated efforts to code the decedent's industry and occupation on death certificates.[3]

We report the results of a pilot population-based surveillance project performed in the Washington, DC metropolitan area, which includes the nation's capital and surrounding areas of Maryland and Virginia. The estimated population of this metropolitan area is 3.25 million persons, 28% of whom are Black.[4] The District itself is home to 639 000 residents, 70% of whom are Black and 19% of whom live below the poverty level. The entire metropolitan area includes 1.5 million workers, 36% of whom work in the District,[5] while the District population itself includes 295 100 workers, 70% of whom work within its borders.[5]

Methods

A manual review of all death certificates maintained by the District of Columbia's Division of Vital Statistics and Research for the period 1980 to 1987

Ann Cottrell is with the Howard University Medical School; Eugene Schwartz and Vincent Kofie are with the Commission of Public Health; and Rosemary Sokas and Laura Welch are with the Division of Occupational and Environmental Medicine at George Washington University Hospital, all in Washington, DC.

Requests for reprints should be sent to Rosemary Sokas, MD, MOH, Division of Occupational and Environmental Medicine, George Washington University Hospital, 2300 K Street, NW, Washington, DC 20037.

This paper was submitted to the journal December 17, 1990, and accepted with revisions May 7, 1991.

TABLE 1—Sentinel Occupationally Related Deaths by Race, Sex, and Age at Death: District of Columbia, 1980–1987 Automated and Manual Search of Death Records[a]

Cause of Death	Race-Sex Group	Number of Cases	Mean Age, y	Range, y
Mesothelioma	Black males	11	67	
	White males	11	67	
	Black females	3	73	
	White females	9	63	
		Group total = 34 (9)	66	34–88
Asbestosis	Black males	3	62	
	White females	1	84	
		Group total = 4 (1)	69	53–84
Silicosis	Black males	3	59	
	White males	3	80	
		Group total = 6 (3)	67	40–80
Anthracosis	Black males	1	71	
	White males	3	85	
		Group total = 4 (2)	82	71–100
Total Number		48 (15)	68	34–100

[a]Numbers in parentheses denote the number of cases identified by automated search.

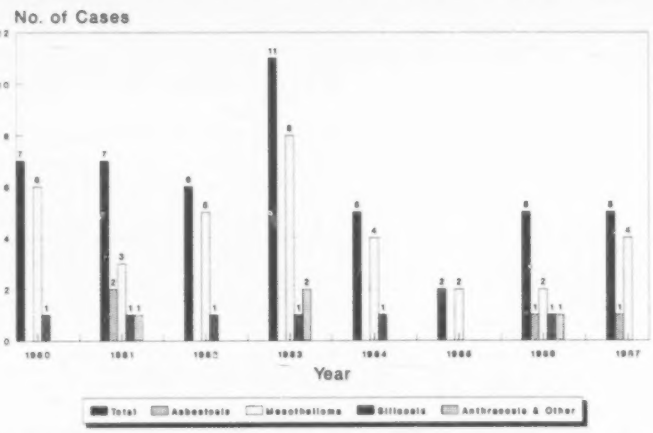


FIGURE 1—Trends in selected sentinel deaths in the District of Columbia, 1980–1987.

(n = 70 537) was performed to identify certificates that had any mention of silicosis,[3] asbestosis,[1] coal worker's pneumoconiosis,[2] or mesothelioma or cancer of the pleura.[9] These four health events that were occupationally related (SHE-Os) (Sentinel Health Event—Occupational) were considered to be almost certainly attributed to occupational or environmental exposure. Additionally, a computerized search using codes from the *International Classification of Diseases,* 9th revision *(ICD-9),*[6] was made for all certificates on which the underlying cause of death was listed as *ICD-9* 163.9 (cancer of the pleura), 500 (coal worker's pneumoconiosis), 501 (asbestosis), 502 (silicosis), or 505 (pneumoconiosis—unspecified). Abstracted data included the decedent's age, sex, race, cause of death, usual place of residence, and usual occupation. Statistical testing of mean ages at death for various race and sex groupings was performed using the *t* test.

Results

The manual search identified 48 SHE-Os, with an average age at death of 68 years (range 34 to 100 years) (Table 1). Males accounted for 35 of the deaths (73%); females accounted for 13 of the deaths (27%), with an average age at death of 69 and 67 years respectively. When delineated by race and sex, Black males accounted for 18 of the deaths (38%), with an average age at death of 65 years. Seventeen deaths (35%) occurred among White males, with an average age at death of 72 years ($t = 1.65$, $P = 0.11$).

Figure 1 describes the number of sentinel deaths that occurred each year, delineated by cause of death. Mesothelioma was recorded on 34 death certificates, with an average age at death of 66 years (range 34 to 88). The occurrence of mesothelioma deaths fluctuated during the period, peaking in 1983. Five of the occupations listed on these death certificates were in the construction trades. Other occupational statements (e.g., housewife, cab driver) may indicate exposure earlier in the decedent's career, including military service, summer or part-time jobs, "carry-home" exposure, or environmental exposure. Two mesothelioma deaths occurred among barbers.

Asbestosis was listed as the cause of death on four death certificates, and silicosis was recorded on six. Five of the occupations listed on these death certificates were in either construction or mining. Anthracosis, or coal worker's pneumoconiosis, was recorded as the cause of death on four death certificates. Only one of the decedents was noted to be a miner.

The computerized search using the *ICD-9* codes listed above identified 15 deaths for the period 1980 to 1987. This number was substantially fewer than the 48 found in the manual search for the same period.

Discussion

Death certificates offer a readily accessible uniform data source for surveillance of occupationally related diseases. The limitations of death certificates are well known and include misclassification

of cause-of-death data as well as often sketchy or incomplete occupational history statements.[7,8] The usefulness of death records is further limited when only a singular underlying cause of death is coded and computerized. The lack of access to computerized multiple cause-of-death files precludes the ability of automated searches to identify cases in which sentinel occupationally related conditions may have been a contributing cause of death. In this study, the computerized review of death certificates coded with a singular underlying cause of death missed over two thirds of the SHE-Os identified by a manual search that considered all listed contributing causes of death.

In addition, conditions such as intrathoracic mesotheliomas are often difficult to identify through death certificates because they are not coded with sufficient morphological information to delineate them from lung cancers or cancers of the pleura. Peritoneal mesotheliomas are harder to identify through death certificates because they may be coded as carcinomatosis, disseminated cancer, or other nonspecific malignancies.[9] Finally, the method by which certifiers complete death certificates compounds the difficulty in identifying occupationally related deaths.[7,10]

The actual number of occupationally caused deaths is likely to be much higher than even that determined by the manual review. We limited our analysis to four sentinel causes of death that are exclusively occupational or environmental in origin. We did not include deaths due to other sentinel occupationally related cancers, such as lung cancer, bladder cancer, oropharyngeal cancer, gastrointestinal cancer, and leukemia, for which nonoccupational contributing factors may predominate. For example, estimates based on industrially exposed cohorts suggest that, for every mesothelioma death, there may be many additional asbestosis cases and asbestos-related lung cancer deaths.[11,12]

Several interesting points may be raised concerning the distribution of work-related deaths in the metropolitan Washington area. First, Blacks accounted for 44% of the deaths while accounting for only 24% of the work force in the metropolitan area and 36% of the work force in the District.[5] Also, the average age at death for Black males tended to be lower than that of White males, suggesting that Blacks in the Washington metropolitan area experience increased exposure in the workplace, face decreased access to health care once symptoms arise, or have a disproportionate number in occupations with the potential for exposure to occupational hazards.

Second, women accounted for more than 25% of all asbestos-related deaths, including one of the four asbestosis deaths. These points are of particular concern, given that death certificate coding is less accurate for minorities than for Whites[13] and that less attention is paid to obtaining occupational histories from female patients than from male patients.[14]

The third finding of interest is that construction workers represented 15% of all SHE-Os while representing only about 8% of the metropolitan area work force. Nationwide, the construction industry ranks third in the rate of work-related fatalities, surpassed by only mining and agriculture. Despite potential risks from health and safety hazards that are commonly encountered in the building and construction trades, the industry remains poorly regulated by the Occupational Safety and Health Administration and is often exempt from many federal standards.

The Washington metropolitan area work force has traditionally been considered "nonindustrial" because of the large proportion of governmental and office workers; hence, it has been deemed free from industrially related diseases. The findings of this pilot project suggest that work-related diseases do occur in areas with a relative absence of what is generally considered to be "heavy industry." Further, these data suggest the need for increased prevention and control efforts, particularly among the construction trades in the District of Columbia metropolitan area. □

References

1. Baker EL, Honchar PA, Fine LJ. Surveillance in occupational illness and injury: concepts and content. *Am J Public Health.* 1989;79(suppl):9–11.
2. Rutstein DD, Mullan RJ, Frazier TM, Halperin WE, Melius JM, Sestito JP. Sentinel health events (occupational): a basis for physician recognition and public health surveillance. *Am J Public Health.* 1983;73(9): 1054–1062.
3. Melius JM, Sestito JP, Seiigman PJ. Occupational disease surveillance with existing data sources. *Am J Public Health.* 1989;79(suppl):46–52.
4. US Dept of Commerce, Bureau of Census. *County and City Data Book–1983: A Statistical Abstract* (suppl). 10th ed. Washington, DC: 1983.
5. US Dept of Commerce, Bureau of the Census. *1980 Census of Population and Housing, Census Tracts: Washington, DC-MD-VA,* Sections 1 and 2. July 1983.
6. World Health Organization. *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death.* 9th rev. Geneva, Switzerland: World Health Organization; 1975.
7. Percy C, Stanek E, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. *Am J Public Health.* 1981;71:242–250.
8. Swanson GM, Schwartz AG, Burrows RW. An assessment of occupation and industry data from death certificates and hospital medical records for population-based cancer surveillance. *Am J Public Health.* 1984;74:464–467.
9. Antman K, Aisner J. *Asbestos-Related Malignancies.* New York, NY: Grune and Stratton; 1987:33.
10. Kircher T, Anderson RE. Cause of death, proper completion of the death certificate. *JAMA.* 1987;258:349–352.
11. Lilienfeld DE, Mandel JS, Coin P, Schuman LM. Projection of asbestos-related diseases in the United States, 1985–2009, I: cancer. *Br J Ind Med.* 1988;45:283–291.
12. Selikoff IJ, Hammond EC, Seidman H. *Mortality Experience of Insulation Workers in the United States and Canada.* New York, NY: Academy of Science; 1979; 330.
13. Becker TM, Wiggins CL, Key CR, Samet JM. Symptoms, signs, and ill-defined conditions: a leading cause of death among minorities. *Am J Epidemiol.* 1990;131(4):664–668.
14. Sokas RK, Diserens D, Johnston MA. Integrating occupational health into the internal medicine clerkship using problem-based learning. *J Gen Intern Med.* 1991;6:450–454.

Maternal Recall Error of Child Vaccination Status in a Developing Nation

Joseph J. Valadez, PhD, SD, and Leisa H. Weld, PhD

ABSTRACT

In the absence of vaccination card data, Expanded Program on Immunization (EPI) managers sometimes ask mothers for their children's vaccination histories. The magnitude of maternal recall error and its potential impact on public health policy has not been investigated. In this study of 1171 Costa Rican mothers, we compare mothers' recall with vaccination card data for their children younger than 3 years. Analyses of vaccination coverage distributions constructed with recall and vaccination-card data show that recall can be used to estimate population coverage. Although the two data sources are correlated (r = .71), the magnitude of their difference can affect the identification of the vaccination status of an individual child. Maternal recall error was greater than two doses 14% of the time. This error is negatively correlated with the number of doses recorded on the vaccination card (r = −.61) and is weakly correlated with the child's age (r = −.35). Mothers tended to remember accurately the vaccination status of children younger than 6 months, but with older children, the larger the number of doses actually received, the more the mother underestimated the number of doses. No other variables explained recall error. Therefore, reliance on maternal recall could lead to revaccinating children who are already protected, leaving at risk those most vulnerable to vaccine-preventable diseases. (*Am J Public Health.* 1991;82:120–123)

Introduction

This paper investigates the validity of mother's recall of a child's vaccination status. Vaccination cards are used as the accurate count of vaccinations to determine the direction and magnitude of maternal recall error. Previous research carried out on vaccination recall focused on American parents remembering their own vaccinations, rather than those of their children.[1] In this paper we recommend against using mother's recall for EPI monitoring and point out the importance of workable vaccination card programs.

Methods

During early 1987, a representative sample of 1680 households with children younger than 3 years was taken in Costa Rica in order to measure health services coverage.[2] Trained interviewers used a pretested survey instrument to collect data. Costa Rica's successful vaccination card policy made feasible the use of these data for a comparison between recall and recorded data.

One child younger than 3 years was randomly selected from the household. The mother was asked to remember the number of doses of BCG, DPT, polio, and measles vaccine that child had received. (The interviewer used body sites and local names for drugs.) Afterward, the vaccination card was requested and vaccination dates were recorded. Up to 11 vaccinations were included: 1 BCG, 4 DPT, 4 polio, and 2 measles.

Recall error was measured as the difference between the number of remembered vaccinations and the number of recorded vaccinations. Positive numbers indicated overestimation, which places children at risk by resulting in community health workers' failing to vaccinate children who still require additional doses. Underestimates result in the waste of resources by vaccinating children with unnecessary doses.

A total of 509 mothers were excluded from the study because they had lost the vaccination card, vaccination dates for any dose were illegible or missing, or the mother refused to try to remember a child's status for any vaccine. Few differences were shown by t tests in which the 1171 included mothers and the 118 excluded mothers with recall data were compared on various demographic variables.

Children who had never had a vaccination card were judged to have never been vaccinated, since Ministry of Health norms require a child to have or be given a vaccination card any time a vaccination is given.

Results

Vaccination Distributions

Distributions of the number of doses from the vaccination cards and from maternal recall are similar (Spearman's r = .65; see Figure 1). This suggests that in the absence of vaccination cards, maternal recall can be used as a valid estimate of coverage of the EPI program in Costa Rica. Both sources of data indicate a successful program. A conclusion of policymakers in this circumstance could be to continue the existing EPI program. However, the important public health question remains: Can maternal recall be used to identify an individual child at risk for vaccine-preventable disease?

Child Vaccination Status

The two data sources are correlated on an individual level (Spearman's r = .71) and the average difference is small, but large differences occur for some children (Table 1). Fourteen percent of the mothers mis-estimated by three or more doses. Thus, although maternal recall can be a

At the time of this study, the authors were with the Harvard Institute for International Development.

Requests for reprints should be sent to Leisa H. Weld, PhD, Gattenhoferweg 7, 6370 Oberursel, Germany.

This paper was submitted to the journal October 1, 1990, and accepted with revisions May 7, 1991.

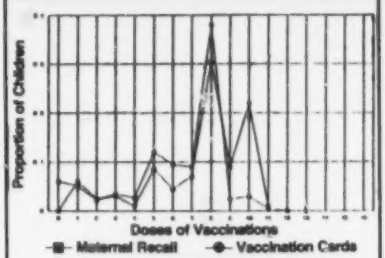


FIGURE 1—Distribution of vaccination coverage using recall and vaccination card data. Average maternal recall = 6.97 doses; Average actual number of doses = 6.50.

TABLE 1—Summary of Differences between Number of Doses Received and Maternal Recall of Vaccinations (n = 1171)

	Percentage of Mothers	Maximum Recall Error
Overestimates	17	10 doses
Underestimates	37	8 doses

Note. Average recall error = −.47 doses.

TABLE 2—Correlation Coefficients of Maternal Error about Child's Vaccination History and Covariates (n = 1171)

	Mother's Memory Error
Number of doses from vaccination card	−0.61*
Mother's recall of doses	0.03
Annual community health worker visits	0.01
Traveling time from home to health center	0.02
Family recorded in health center records	−0.05
Child recorded in health center records	0.04
Number of persons living in house	0.08
Number of adults in house	−0.03
Number of adults in house who read and write	−0.03
Number of educated adults in house	−0.01
Mother's age	−0.01
Mother's education	−0.05
Number of children living	0.11
Age of child at time of interview	−0.35*
Months of residence in household	−0.05
Socioeconomic status	−0.04

*$P < .0001$.

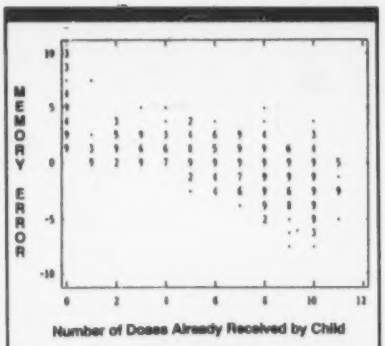


FIGURE 2—Relationship between maternal recall error and number of doses received by children as recorded on the vaccination card. 1 = 1 observation, 2 = 2 observations, . . ., · = 10 or more observations.

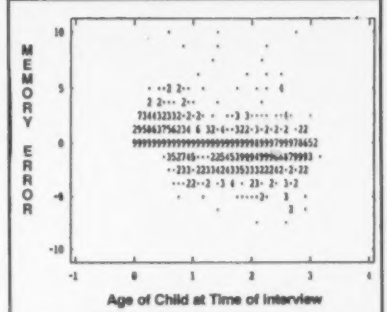


FIGURE 3—Relationship between maternal recall error and child's age. 1 = 1 observation, 2 = 2 observations, . . ., · = 10 or more observations.

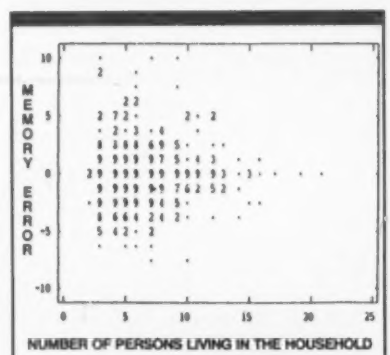


FIGURE 4—Relationship between maternal recall error and number of persons living in the household. 1 = 1 observation, 2 = 2 observations, . . ., · = 10 or more observations.

valid estimate of coverage, it should not be used to determine an individual child's vaccination status.

Spearman's correlations between maternal recall error and potential explanatory variables are given in Table 2. Number of doses on the vaccination card and the child's age showed large negative correlations with recall error (Figures 2 and 3). The correlation with age is primarily an artifact, because mothers with very young children who received few doses cannot underestimate the number of doses. The negative correlation between recall error and actual doses has important implications. It suggests that mothers whose children are most at risk for vaccine-preventable diseases tend to overestimate their children's vaccinations. Mothers whose children are the least vulnerable tend to exhibit conservative preventive behavior by underestimating their children's coverage.

Further analyses show that the range of recall errors decreases as the number of contacts with community health workers increases (e.g., number of previous children, number of persons living in the household; see Figure 4). However, there is no relationship between frequency of contact with community health workers and the likelihood of overestimates relative to underestimates of vaccination status, and so frequency of contact with the health system does not explain recall error. Other variables exhibited no significant effect.

Reanalysis of the data excluding the 62 mothers who claimed never to have had a vaccination card for their child did not change the results.

Type of Vaccination Regime

Cross-tabulations of maternal recall error against number of doses received are reported in Table 3 for each vaccination type. As with the overall results, the recall errors for measles, polio, and DPT vaccinations were inversely related to the actual number of vaccination doses. All of the mothers believed their children had received a BCG vaccination. Mothers whose children had received the vaccination remembered, but mothers whose children had not received it overestimated the number of doses.

TABLE 3—Maternal Recall Error by Number of Doses Received for Four Vaccination Regimes

Error	Number of Doses 4	3	2	1	0	Total
			DPT			
4	0	0	0	0	1	1
3	0	1	0	0	9	10
2	1	1	1	3	11	17
1	5	11	7	10	11	44
0	86	365	106	83	85	725
−1	166	93	23	2		284
−2	32	41	2			75
−3	10	2				12
−4	3					3
Total	303	514	139	98	117	1171
			Polio			
4	0	0	0	0	1	1
3	0	0	0	0	7	7
2	0	0	0	1	9	10
1	4	4	11	7	14	40
0	75	369	110	87	87	728
−1	164	105	22	3		294
−2	37	35	0			72
−3	14	2				16
−4	3					3
Total	297	515	143	98	118	1171
			Measles			
4			0	1	0	1
3			0	1	0	1
2			0	8	5	13
1			1	36	57	94
0			24	623	293	940
−1			65	55		120
−2			2			2
Total			92	724	355	1171
			BCG			
2				0	43	43
1				25	124	149
0				979	0	979
Total				1004	167	1171

Note. DPT $\chi^2 = 350.4$, $P < .001$; polio $\chi^2 = 385.4$, $P < .001$; measles $\chi^2 = 303.3$, $P < .001$; BCG $\chi^2 = 1000.9$, $P < .001$.

Discussion

A comparison of the number of vaccinations recorded in vaccination cards and the number recalled by mothers revealed important recall errors. Maternal recall error is correlated with vaccination frequency of the child. Whatever the actual number of doses, the recalled number of doses regresses toward the population average. The greater the frequency of contacts with the health system, the smaller the range of recall errors.

Although no data reported here explain the negative correlation, further research should consider memory problems and the effect of social norms regarding vaccinations.

The results reported above suggest that relying on maternal recall of a child's vaccination history may result in erroneous decisions about whether to vaccinate the child. Children who are most at risk for vaccine-preventable diseases could be the most likely to remain unvaccinated, owing to their mothers' overestimation of doses received. Children who are least vulnerable, because of prior vaccination, are most likely to receive redundant vaccinations, owing to their mothers' underestimation. Nonetheless, maternal recall rather than vaccination cards can be used to estimate coverage.

As a result, EPI policies using recall data may have limited construct validity and thus ought to support the development of national vaccination card programs. Such a program's cost could potentially be offset by the savings resulting from not vaccinating children who already are up to date in their vaccinations. □

Acknowledgments

The work upon which this study is based was performed under a subagreement with the Center for Human Services under its Cooperative Agreement NO. DPE-5920-00-A-5056-00 with the US Agency for International Development. We would like to express our gratitude to the coordinators of the field staff in Costa Rica, William Vargas Vargas and Silvia Boada Martinez; to the Director General of Health of Costa Rica at the time of the study, Dr Carlos Valerín; to the Country Representative for the Pan American Health Organization, Raul Pena; and to Dr James Heiby, the Program Officer of the PRICOR Project. Without their support this research would not have been possible.

References

1. Le CT. Parental knowledge of their own immunization of poliomyelitis. *JAMA.* 1985; 254:608–609.
2. Valadez JJ. *Assessing Child Survival Strategies: A Field Test in a Developing Nation.* Cambridge, Mass: Harvard University Press; 1991.

AZC23-1

American Journal of Public Health

February 1992, Vol. 82, No. 2 Established 1911

86389 028
ED KIANKA
CONGRESSIONAL INFORMATION SVC
4520 EAST-WEST HIGHWAY
BETHESDA MD 20814

American Public Health Association

Featuring Women's Health

Editorials

165 Women Health Workers: Past and Present *E. Fee and R. R. Korstad*

166 Men Who Have Sex with Men: Continued Challenges for Preventing HIV Infection and AIDS *A. R. Lifson*

Featuring Women's Health

168 Prevention: The Cost-effectiveness of the California Diabetes and Pregnancy Program *R. M. Scheffler, L. B. Feuchtbaum, and C. S. Phibbs*

176 Maternal Smoking and the Risk of Polyhydramnios *W. Myhra, M. Davis, B. A. Mueller, and D. Hickok*

180 A Cost-effectiveness Analysis of Three Staffing Models for the Delivery of Low-Risk Prenatal Care *E. A. Graveley and J. H. Littlefield*

185 Medicaid Prenatal Care: A Comparison of Use and Outcomes in Fee-for-Service and Managed Care *J. W. Krieger, F. A. Connell, and J. P. LoGerfo*

191 Attitudes about Infertility Interventions among Fertile and Infertile Couples *L. J. Halman, A. Abbey, and F. M. Andrews*

195 Nonclinical Factors Associated with Surgery Received for Treatment of Early-Stage Breast Cancer *E. R. Satariano, G. M. Swanson, and P. P. Moll*

199 Breast Cancer Screening Practices among Users of County-Funded Health Centers vs Women in the Entire Community *D. S. Lane, A. P. Polednak, and M. A. Burg*

204 Geographic and Socioeconomic Variation in the Onset of Decline of Coronary Heart Disease Mortality in White Women *S. Wing, E. Barnett, M. Casper, and H. A. Tyroler*

210 Vaginal Douching among Women of Reproductive Age in the United States: 1988 *S. O. Aral, W. D. Mosher, and W. Cates*

215 Employment Status and Heart Disease Risk Factors in Middle-Aged Women: The Rancho Bernardo Study *D. Kritz-Silverstein, D. L. Wingard, and E. Barrett-Connor*

249 Cancer Anxiety and Attitudes toward Mammography among Screening Attenders, Nonattenders, and Women Never Invited *I. T. Gram and S. E. Slenker*

252 Medicaid Coverage of Screening Tests for Breast and Cervical Cancer *L. P. Boss and F. H. Guckes*

254 Distribution of HIV Type 1 Infection in Childbearing Women in California *F. J. Capell, D. J. Vugia, V. L. Mordaunt, W. D. Marelich, M. S. Ascher, A. I. Trachtenberg, G. C. Cunningham. S. S. Arnon, and K. W. Kizer*

257 Cesarean Section Rates in Italy by Hospital Payment Mode: An Analysis Based on Birth Certificates *R. Bertollini, D. DiLallo, T. Spadea, and C. Perucci*

262 Professional Relationships between Midwives and Physicians: Collaboration or Conflict? *L-M. Baldwin, H. L. Hutchinson, and R. A. Rosenblatt*

264 Factors Associated with Artificial Feeding in Shanghai *Y. Chen*

267 Smoking-Related Behavior, Beliefs, and Social Environment of Young Black Women in Subsidized Public Housing in Chicago *C. Manfredi, L. Lacey, R. Warnecke, and M. Buis*

(continued page 153)

Contents

American Journal of Public Health

"Information for Authors" appeared on page 138 of the January 1992 issue and may also be obtained from the Journal Editorial Office. An abbreviated version is printed in this issue.

American Journal of Public Health
1015 Fifteenth Street, NW
Washington, DC 20005

APHA
American Public Health Association

February 1992, Vol. 82, No. 2 ISSN: 0090-0036 Established 1911

Articles

220 AIDS in Bisexual Men in the United States: Epidemiology and Transmission to Women *S. Y. Chu, T. A. Peterman, L. S. Doll, J. W. Buehler, and J. W. Curran*

225 Knowledge and Attitudes about AIDS among Corporate and Public Service Employees *J. K. Barr, J. M. Waring, and L. J. Warshaw*

229 Workplace Compliance with a No-Smoking Law: A Randomized Community Intervention Trial *N. A. Rigotti, D. Bourne, A. Rosen, J. A. Locke, and T. C. Schelling*

236 Attitudes Toward Genetic Testing of Amish, Mennonite, and Hutterite Families with Cystic Fibrosis *S. R. Miller and R. H. Schwartz*

243 The Accuracy of Medicare's Hospital Claims Data: Progress Has Been Made, but Problems Remain *E. S. Fisher, F. S. Whaley, W. M. Krushat, D. J. Malenka, C. Fleming, J. A. Baron, and D. C. Hsia*

Public Health Briefs

273 Smoking Initiation and Cessation in Relation to Body Fat Distribution Based on Data from a Study of Swedish Women *L. Lissner, C. Bengtsson, L. Lapidus, and C. Björkelund*

275 Surveillance of Occupational Lead Exposure in New Jersey: 1986 to 1989 *A. Tepper*

278 Inaccuracies in Certification of Nonmelanoma Skin Cancer Deaths *M. A. Weinstock, H. A. Bogaars, M. Ashley, V. Litle, E. Bilodeau, and S. Kimmel*

281 Provider Compliance with Recommended Dietary Fluoride Supplement Protocol *S. M. Levy and G. Muchow*

284 Condom Use in Multi-Ethnic Neighborhoods of San Francisco: The Population-Based AMEN (AIDS in Multi-Ethnic Neighborhoods) Study *J. A. Catania, T. J. Coates, S. Kegeles, M. T. Fullilove, J. Peterson, B. Marin, D. Siegel, and S. Hulley*

Public Health Then and Now

288 Women Healers of the Middle Ages: Selected Aspects of Their History *W. L. Minkowski*

Public Health and the Law

296 Mum's the Word: The Supreme Court and Family Planning *W. K. Mariner*

Other Journal Departments

154 Masthead
157 Information for Authors
158 Advertisers' Index
302 Notes from the Field
305 Book Corner
309 Association News/Call for Abstracts
319 Job Opportunities

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.

APHA OFFICERS

Joyce C. Lashof, MD, *President*
C. William Keck, MD, MPH, *Immediate Past President*
Helen Rodriguez-Trias, MD, *President-Elect*
Arnold E. Greenberg, SM, *Vice-President (USA)*
Trevor Hancock, MB, MHSc, *Vice-President (Canada)*
Julio Teja Perez, MD, *Vice-President (Latin America)*
Caswell A. Evans, Jr, DDS, MPH, *Treasurer*
Frances M. Veverka, RS, MPH, *Chair, Executive Board*
George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*
William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD

Frances M. Veverka, RS, MPH (1992), *Chair*
Marilyn Aguirre-Molina, EdD, MS (1993)
Brian A. Dobrow, MPH (1992)
Caswell A. Evans, Jr, DDS, MPH, *Treasurer*
Kristine M. Gebbie, BSN, MN (1994), *Vice-Chair*
Alan R. Hinman, MD, MPH (1995)
Janet Heinrich, DrPH, MPH (1993)
Margaret H. Jordan, RN, MPH (1992)
C. William Keck, MD, MPH, *Immediate Past President*
Margo D. Kerrigan, MPH (1995)
Joyce C. Lashof, MD, *President*
Barry S. Levy, MD, MPH (1993)
Helen Rodriguez-Trias, MD, *President-Elect*
Allan G. Rosenfield, MD (1994)
Reed V. Tuckson, MD (1994)
Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO

Willam H. McBeath, MD, MPH, *Executive Director*
Joanne E. Lukomnik, MD, *Chair, Action Board*
George Pickett, MD, MPH, *Chair, Program Development Board*
George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc., 1015 Fifteenth St., NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50,000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health,* 1015 Fifteenth St., NW, Washington, DC 20005.

Information on Manuscript Preparation is published in this Journal in January. The Journal publishes reports of original research, demonstrations, evaluations, and other articles covering current aspects of public health. Manuscripts are accepted for consideration with the understanding that they, or their essential substance, have not been published previously or accepted for publication and are not under consideration elsewhere.

Subscription price, individual, is $80/year US, $120/year international; **institutional,** $160/year US, $240/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to American Journal of Public Health, 1015 Fifteenth St., NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact: Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact: Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols. 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS Bibliographic Retrieval Service Inc, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc., Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Prevention: The Cost-effectiveness of the California Diabetes and Pregnancy Program

Richard M. Scheffler, PhD, Lisa B. Feuchtbaum, MPH, and Ciaran S. Phibbs, PhD

ABSTRACT

Background. The California Diabetes and Pregnancy Program is a new preventive approach to improving pregnancy outcomes through intensive diabetes management preconception and early in pregnancy.

Methods. Hospital charges and length of stay data were collected on 102 program enrollees and 218 control cases. Ninety program enrollees and 90 control cases were matched on mother's age, White's classification, and race. Regression models controlled for these variables in addition to MediCal status, birth weight, and enrollment in the program.

Results. Hospital charges were about 30% less for program participants and days in the hospital were roughly 25% less. The program effects were larger for women that enrolled before 8 weeks gestation. More serious diabetics were also found to have larger reductions in charges and days.

Conclusion. After adjusting for inflation and differences in charges across hospitals, $5.19 is saved for every dollar spent on the program. (*Am J Public Health.* 1992;82:168–175)

Introduction

The cost savings potential of many prevention programs may not be realized for years, making these programs unattractive investments for health policy decision makers.[1,2] Fortunately, many studies of the cost savings potential of prenatal prevention programs need only look at the immediate future (i.e., 1-year care) to assess the cost effectiveness of the intervention. These studies have generally focused on the benefits of prenatal care or the prevention of low birth weight. Importantly, in the past 10 years there has been increasing recognition that the infant and maternal complications associated with diabetes in pregnancy can be prevented.[3–7] This paper provides the first cost-effectiveness analysis of a new and widely accepted prenatal prevention program, the California Diabetes and Pregnancy Program (CDAPP).

The CDAPP, also known as the "Sweet Success" program, focuses on improving maternal and infant perinatal outcomes in pregnancies complicated by diabetes. The program's multidisciplinary team approach to care emphasizes early recruitment of pre-pregnant (prior to conception, or "pre-conception") and pregnant women with diabetes. Once enrolled, the women receive comprehensive prenatal care, including nutrition, education, and support services that have been specifically developed to meet the needs of diabetic women during pregnancy. This management approach has been shown to reduce diabetes-related complications during pregnancy to the rate seen in the nondiabetic population.[8–10] This approach to care, which has been shown to be effective in improving birth weight in the nondiabetic population,[11] is particularly well suited for managing high-risk pregnancies.

Although maternal and infant mortality have been greatly reduced by advances in insulin therapy over the past 10 years, preexisting diabetes mellitus remains a significant risk to the unborn child. Congenital malformations still account for nearly 50% of the deaths among infants born to diabetic mothers today.[12] Serious malformations are three to four times more likely to occur among infants of diabetic women than among infants of nondiabetics,[13] and affect 6% to 9% of all diabetic pregnancies.[3] These malformations occur very early in pregnancy, at a time when the diabetic woman may not yet be aware that she is pregnant.[14] For these malformations to be prevented, it is critical that maximum diabetes control be maintained prior to pregnancy and in the early weeks of embryogenesis.[15]

It is still not clear what specific biological mechanism accounts for these malformations, but it is widely accepted that the congenital malformation rate is correlated with the degree of glycemic control during pregnancy.[7,16] Although the appropriate degree of control is still being de-

Richard M. Scheffler is with the Schools of Public Health and Public Policy and Lisa B. Feuchtbaum is a doctoral candidate with the School of Public Health, both at the University of California, Berkeley. Ciaran S. Phibbs was with the School of Public Health, Columbia University, New York, NY, at the time of this study, and is now with the Center for Health Care Evaluation, Veterans Administration Medical Center, Palo Alto, Calif.

Requests for reprints should be sent to Richard M. Scheffler, PhD, School of Public Health, Warren Hall, University of California, Berkeley, Berkeley, CA 94702.

This paper was submitted to the journal February 2, 1991, and accepted with revisions September 13, 1991.

bated,[17] the more tightly diabetes is controlled before and during pregnancy, the better the pregnancy prognosis. In studies in which women have achieved and maintained normal blood glucose levels before and throughout pregnancy, the congenital malformation rate has been reduced to the rate seen in the nondiabetic population.[3,18,19]

The aim of the present study was to determine the cost-effectiveness of the CDAPP. We compared the hospital charges and lengths of stay for mothers and their infants in the CDAPP with those for a sample of births to diabetic mothers not in the program. A regression framework was used to estimate the differences in hospital charges and lengths of stay between the CDAPP group and the control group; the regression models included measures of age, race, diabetes duration and severity, and the birth weight of the child.

Methods

Study Design

We used a case-control study design and incorporated the costs of the program into the cost-effectiveness calculations. For ethical and practical reasons, we did not randomly assign women to the study group or the control group. However, we minimized the differences between cases and controls by matching participants on maternal age, ethnicity, and duration and severity of diabetes. To minimize the effect of selection bias, we chose the control hospitals from areas outside the catchment areas of the CDAPP hospitals.

Hospital Site Selection

Data for the 102 CDAPP cases were collected for all program participants at three CDAPP hospitals (Children's Hospital of San Francisco, n = 18; Sutter Memorial Hospital, n = 36; and Loma Linda University Medical Center, n = 48). We included all births to program participants between July 1, 1986, and July 30, 1988, for whom complete data could be obtained. All obstetricians who regularly delivered infants at these hospitals referred their diabetic cases to the program. Thus, virtually all of the diabetic deliveries at these hospitals were included in our data. These hospitals were selected because they were the first to implement the CDAPP and their programs were well established at the time data collection started. All CDAPP hospitals have a level 3 neonatal intensive care unit.

Data for the 218 control cases were collected from diabetic deliveries identified at five hospital sites throughout the state: Alta Bates Hospital in Berkeley (n = 89; 1984–1987), Huntington Memorial Medical Center in Pasadena (n = 71; 1985–1987), Sharp Memorial Medical Center in San Diego (n = 12; 1985–1987), St. Agnes Medical Center in Fresno (n = 15; 1985–1986), and Fresno Community Hospital (n = 31; 1985–1986). The control hospitals were selected from among hospitals that (1) were in regions that had yet to have a CDAPP implemented, (2) had at least 2000 births per year, and (3) had at least a level 2 neonatal intensive care unit.

The demographic characteristics and health status indicators of mothers and babies at these hospitals were similar. Information from the 1986 Maternal and Child Health Data Base Statistical Appendix[20] was used to compare demographic characteristics of study and control hospitals. For study and control hospitals, respectively, the average number of births in 1986 was 3867 and 4355; Black births represented 10.4% and 14.5% of total births; the percentage of women enrolled in prenatal care in the first trimester of pregnancy was 85% and 81%; the observed mortality rate was 1.8% and 1.2%; the perinatal complication rate was 17.5% and 13.8%; and the congenital malformation rate was .70% and .56%.

Only women with overt diabetes were included in the study. The severity and duration of diabetes were assessed with White's classification system[21]:

- A Asymptomatic diabetes recognized prior to pregnancy; chemical diabetes.
- B Disease that began after age 20, or that has been present for 9 or fewer years with no vascular disease.
- C Disease that began between the ages of 10 and 19, or that has been present for between 10 and 19 years with no vascular disease.
- D Disease that began before age 10, or that has been present for 20 or more years, or with vascular disease, such as background retinopathy or calcified iliac or femoral arteries.
- F Disease complicated by diabetic nephrology.
- H Disease complicated by cardiovascular disease.
- R Disease complicated by proliferating retinopathy.
- T Disease complicated by renal transplant.

Both CDAPP cases and control cases were excluded from the study if their White's classification could not be identified from the patient's medical record. Because of the very small number of cases in classes F, H, R, and T, these classes were combined with class D in a group we labeled D+.

CDAPP cases were identified by tracking enrollment records at each of the three CDAPP collection sites. Clinical and demographic data for each case were obtained from the CDAPP program database. Hospitalization charge data were collected for each mother and baby after delivery by means of the billing summary reports obtained from the hospital records. Complete billing histories were obtained for each mother for the duration of pregnancy and delivery, and for the babies from birth to discharge.

Control cases were identified by searching the computerized medical record systems for the specific ICD-9 code specifying diabetic deliveries. Billing records were then requested for the mothers and their babies to reflect all inpatient hospital admissions of the mother during her pregnancy and delivery, in addition to infant hospital charge data for the period from birth through discharge. In the event that the baby was transferred to another hospital for care, billing summaries were requested for the care delivered at the second institution.

Costs for the care provided by the CDAPP were estimated by means of "time-motion" methodology. The cost of staff time was calculated on the basis of expected staff time necessary to deliver patient care services as specified in the state *Guidelines For Care.*[22] The costs of CDAPP services were estimated on the basis of average salaries, fringe benefits, and overhead. For women who enrolled before conception, these costs averaged $1300. For late enrollees (enrolled after 12 weeks' gestation), the costs averaged $800.

The Match

To control for individual risk factors that may have an ex ante effect on hospital charges, we matched the data on mother's age, race, and severity and duration of the mother's diabetes using White's classification. Race was classified into two groups (Black and non-Black) because Black infants have been shown to be significantly different from infants of other races with respect to incidence of low

birth weight, birth-weight–specific mortality, and overall mortality. Mother's age is usually associated with risk status, especially for older mothers and the infants of younger mothers. We used three age groups (less than 20 years, 20 to 34 years, and 35 and over). Finally, White's classification was stratified into three groups: A, B and C, and D+, yielding a total of 18 possible cells.

Each CDAPP participant was randomly matched with an individual from the control group who was of the same age, ethnicity, and White's class. This procedure yielded a data set of 180 matched cases, 90 cases from the CDAPP group and 90 cases from the control group. We excluded a control group infant who incurred a hospital bill of over $1.3 million (length of stay, over 400 days). Infants requiring hospitalization beyond 6 months of age are quite rare; one was included in our sample just by chance. Including this observation would greatly increase the savings due to the CDAPP.

Variable Construction

The dependent variable is total hospital charges, which was constructed by aggregating the charges for the mothers and the babies. Charges were adjusted using the hospital component of the Consumer Price Index, so that all charges are in June 1988 dollars. Because the study hospitals are in several different labor markets across the state, a second adjustment was required. Data on wages for hospital personnel from the 1986 American Hospital Association Survey of Hospitals were used to construct a hospital-specific wage index. The hospital wage index controls for the difference in labor costs, which are the largest component of hospital costs. The average wage, weighted by the number of cases from each hospital, was used as the base for the index. Hospital-specific charges were multiplied by 1, plus or minus the percentage by which the hospital's wage differed from the weighted average wage across hospitals. Normalizing charges with this wage index makes the charges across hospitals more comparable.

The use of hospital charges to conduct a cost-effectiveness analysis has some limitations; differences in charges can be due to accounting practices or hospital inefficiencies. Length of stay is less likely to be subject to these differences. We therefore conducted a parallel analysis, using length of stay as the dependent variable. Length of stay, like charges, was aggregated for each maternal and infant hospitalization. Because both charges and length of stay have a very skewed distribution, the natural logarithms of charges and length of stay were used as the dependent variables in the regression analysis. Independent variables were included to control for risk factors that independently affect the costs of care and the length of hospital stay. These factors included maternal age, race, diabetes classification, birth weight, and source of payment. Variables were also included to test for the impact of the CDAPP.

Binary variables were used to control for the presence of these risk factors. Almost all of the births were paid for by either private insurance or MediCal (the Medicaid program for California). Therefore, a binary variable for a MediCal admission was included. Duration and severity of diabetes was controlled for with White's classifications. These classifications were aggregated into four groups: A, B, C, and D+ (includes classes F, H, R, and T). Ethnicity was controlled for with a binary variable for race (Black or non-Black). Maternal age was controlled for by including the mother's age, in years, at the time the birth occurred. Birth weight was controlled for with a set of binary variables: less than 1500 g, between 1500 g and 2500 g, and greater than 4000 g. Birth weights between 2500 g and 4000 g were the excluded category. The inclusion of birth weight in the model is appropriate because maternal diabetes tends to affect birth weight.[23,24] Moreover, previous research shows that birth weight is the single best predictor of newborn mortality and costs.[25]

To test for simultaneity bias, we reestimated the models reported below, excluding birth weight as a regressor. We also estimated models with birth weight as the dependent variable. These estimates showed no program effect on birth weight, and showed that excluding birth weight had a minimal effect on the estimated program coefficients. Thus, we are confident that any potential simultaneity bias is small.

We used several different specifications to measure the impact of the CDAPP in our regression model. The basic specification included a binary variable for our CDAPP participants. However, to test the effect of time of enrollment in the CDAPP, we created binary variables for mothers enrolled before 8 weeks' gestation and one for mothers enrolled after 8 weeks' gestation (the study did not contain enough preconception cases to allow a completely separate analysis of these cases).

Structure of the Regression Models

All of the regression models control for mother's age, MediCal status, White's class, race (Black), and birth weight, and include a CDAPP effect variable. Model 1 presents the results of the basic model. Model 2 splits the program effect variable by time of enrollment in CDAPP, the expectation being that those women who enrolled earlier would incur lower costs and length of stay. Model 3 attempts to differentiate the degree of the program effect across severity of diabetes according to White's classification system. To test for severity of diabetes, the program effect variable, a binary variable for the CDAPP cases, is interacted with the White's class variables. This procedure splits the control variable into three mutually exclusive groups by White's class (classes A and B, class C, and class D+). The controls for White's class remain in the model. The excluded category becomes control cases whose diabetes is defined by White's class A or B. Thus, the coefficient for each of these new variables gives an estimate of the differential effect of the program for each White's class group. Given the smaller sample sizes, we expected limited significance with these interaction variables.

Cost-effectiveness Calculations

Since the dependent variables for all of the models are in logarithms, direct interpretations of the charge or length-of-stay differences in the program are nontrivial. The log transformation moves the distribution toward normality, but it is still not normal. When the distribution of errors is not log-normal, the conventional formula for converting logarithms back to levels is inconsistent. Therefore, we used the smearing method developed by Duan[26] to provide consistent estimates.

The savings due to CDAPP do not occur instantly (i.e., costs of the program precede the hospital cost savings at delivery), so they must be discounted to present-value terms. To do this we allocated all hospitalization costs to the time of delivery, using an 8% discount rate. Given that these costs include maternal and infant costs, the time of delivery is probably near the midpoint of the time these costs were incurred. We used two different program patterns for this part of the evaluation. The first used the average time from entry into the program to delivery. This information was obtained from program

records for all mothers who entered the program between January 1, 1986, and December 31, 1987. Average time was calculated to be 26.5 weeks. This time reflects a mixture of pre- and post-conception enrollments. This method evaluates the program as it was actually provided; for notation purposes we refer to this as the Pre/Post analysis.

CDAPP was designed as a pre-conception program, so in our second analysis we looked at the effects of an all–pre-conception program under conservative assumptions. We assumed that the pre-conception program would yield no more benefits than would the existing mix of programs (see method 1 above), and that women in this program would take an average of 6 months to conceive once the go-ahead was given. This 6-month period, plus the estimated 3-month program time, results in an 18-month lag from time of enrollment to time of delivery. We refer to this as the All Pre analysis. To be conservative in our estimate of the program savings, we used an 8% discount rate for the evaluation. This procedure produced discount factors of 0.962 for the Pre/Post analysis and 0.891 for the All Pre analysis. Note that we are using a very restrictive definition of program savings—only the savings in maternal and neonatal costs. Longer-term medical costs due to adverse birth outcomes, which are potential savings of the CDAPP, and any potential health benefits to the mother are not included.

The discounted savings were compared with the program costs. Even though the program treatment costs were spread over time, they were all allocated to the time of initial treatment. Treatment costs of $1300 for pre-conception enrollment and $800 for post-conception enrollment were obtained from a previous analysis of the education component of the program. For the Pre/Post analysis, we used the Pre/Post mix of our sample to calculate a weighted average of $820. These costs were inflated further to reflect the program dropout rate and the failure of some participants to give birth to a live infant. Program data to date show that over 70% of the participants have given birth to a live infant, yielding effective program costs per live infant of $1857 for the All Pre analysis and $1171 for the Pre/Post analysis. These numbers are almost certainly overestimates of the per-delivery program costs, inasmuch as some of the mothers counted in these figures who ultimately will give birth to a live infant have not yet had time to do so.

TABLE 1—Cases and Controls by Selected Demographic and Clinical Characteristics

	All Cases		Matched Cases	
	CDAPP (n = 102)	Control (n = 218)	CDAPP (n = 90)	Control (n = 90)
Mean charge, $				
Total	14 958	18 449	15 344	21 699
Mother	9379	10 521	9426	11 025
Baby	5579	7927	5918	10 674
Mean length of stay, days				
Total	14.7	15.8	15.1	18.0
Mother	9.0	9.5	9.2	10.5
Baby	5.8	6.3	6.0	7.4
White's class, %				
A	1.9	26.1	2.2	2.2
B	41.1	55.5	46.6	66.6
C	27.4	9.6	30.0	10.0
D/F/H/R	29.4	8.7	21.1	21.1
Mean charge by White's class, $				
A	16 000	12 425	16 000	10 450
B	11 018	16 089	11 018	14 022
C	16 951	23 217	17 064	23 474
D/F/H/R	18 572	46 285	22 434	46 285
Mean LOS by White's class, days				
A	14.5	11.1	14.5	10.0
B	10.8	15.0	10.8	13.8
C	16.5	19.6	16.4	18.1
D/F/H/R	18.8	32.1	24.8	32.1
Average maternal age, y	27.7	29.8	28.0	28.8
Birth weight, %				
<1500 g	0.9	1.8	1.1	2.2
1501–2500 g	10.7	7.3	11.1	5.6
>2500–4000 g	71.5	74.3	71.1	77.8
>4000 g	16.7	16.5	16.7	14.4
Race, %				
Black	12.8	21.1	11.1	11.1
Non-Black	87.2	78.9	88.9	88.9
Payor, %				
Non-MediCal	63.7	63.8	62.2	62.2
MediCal	36.3	36.2	37.8	37.8

Results

Selected demographic and clinical characteristics of the cases and controls are presented in Table 1. These characteristics are average charges and length of stay, percentage by White's class, charges and length of stay by White's class, average age of mother, percentage by specific birth-weight group, percentage by race, and percentage by payor. The relative differences between the CDAPP cases and the control group are similar for both the entire sample and the matched sample, except for those factors over which the match was stratified. For the entire sample, the CDAPP group contained women with more severe diabetes; the majority of the women in the control group were of White's classes A and B (26% and 55%, respectively), whereas the majority of the CDAPP group were of classes B, C, and D (41%, 27%, and 29%, respectively). Average length of stay was slightly higher in the control group than in the CDAPP group, for mothers as well as babies. The control group had slightly more babies in the <1500 g birth-weight group, whereas the CDAPP group had more babies in the 1501–2500 g birth-weight group. Control group mothers were slightly older, and more were Black, than mothers in the CDAPP group. The proportions of MediCal and non-MediCal payors were equal in both groups.

When all cases are included, the control group charges (for mother and baby combined) were 23% higher than those of the CDAPP group ($18 449 vs $14 958). For the matched cases, the combined charges of the control group were 41% higher than those of the CDAPP group ($21 699 vs $15 344). Most of this difference can be attributed to charge differences for the babies, particularly in the matched comparison ($10 674 for controls vs $5918 for CDAPP). Charges and length

TABLE 2—Selected Characteristics by Length of Gestation at Program Enrollment: CDAPP Cases

	Before 8 weeks (n = 24)	After 8 weeks (n = 78)
Mean charge, $		
Total	11 199	16 115
Mother	8936	9516
Baby	2263	6599
Mean length of stay, days		
Total	10.9	16.1
Mother	7.1	9.7
Baby	3.9	6.4
Average maternal age, y	30.0	27.1
White class, %		
A	0.0	2.6
B	50.0	38.5
C	25.0	28.2
D+	25.0	30.8
Average birth weight, %		
<1500 g	0.0	1.3
1501–2500 g	12.5	10.3
>2500–4000 g	75.0	58.8
>4000 g	12.5	18.0

TABLE 3—Estimated Coefficients, Regression Model 1

	All Data		Matched Data	
Independent variable	Charges	Length of Stay	Charges	Length of Stay
Intercept	9.1929	2.4817	8.5148	1.8483
	(31.611)	(10.233)	(20.600)	(4.940)
Maternal age	−0.0061	−0.0075	0.0151	0.0111
	(−0.706)	(−1.033)	(1.203)	(0.963)
White's class A	−0.4699	−0.3796	0.0925	−0.0253
	(−3.736)**	(−3.658)**	(0.231)	(−0.072)
White's class C	0.4153	0.2637	0.4750	0.2897
	(2.877)**	(2.177)*	(2.757)**	(1.860)*
White's class D+	0.5448	0.4064	0.7531	0.6245
	(3.718)**	(3.223)**	(4.590)**	(4.089)**
MediCal	0.3087	0.2401	0.2290	0.2586
	(2.997)**	(2.780)**	(1.734)*	(2.133)*
Black	0.2542	0.2627	0.4274	0.3633
	(2.081)*	(2.524)**	(2.214)*	(1.950)*
CDAPP enrollment	−0.3026	−0.2618	−0.2469	−0.1916
	(−2.821)**	(−2.827)**	(−2.068)*	(−1.731)*
Birth weight <1500 g	2.1878	1.6794	2.2986	1.7557
	(5.960)**	(5.551)**	(5.048)**	(4.374)**
Birth weight 1500–2500 g	0.6179	0.4405	0.6278	0.4992
	(3.758)**	(3.140)**	(2.969)**	(2.506)**
Birth weight >4000 g	0.1397	0.1218	0.1385	0.1543
	(1.133)	(1.146)	(0.854)	(0.980)
Adjusted R^2	0.259	0.237	0.270	0.228

Note. Estimated dependent variables are log charges and log length of stay for mothers and babies. *T* statistics are given in parentheses.
*$P \leq .05$ (one-sided).
**$P \leq .01$ (one-sided).

of stay increased with the severity of diabetes. Patient charges were higher when Medicaid was the primary payor, and within this group, the charges for the control cases were always higher than those for the CDAPP cases. Blacks incurred higher average charges than did non-Blacks, across the board. Among non-Blacks, charges were much lower for the CDAPP cases than for the control cases, whereas there was essentially no difference in charges for Blacks.

Table 2 contains charges for CDAPP participants by length of gestation at program enrollment: before 8 weeks (this group includes the pre-conception enrollees) and after 8 weeks. The 8-week cutoff was used because this is the time in organogenesis when major congenital malformations can occur. Lower maternal and infant charges were incurred by the women enrolled earlier in the program (the charges of the after–8-week group were 44% higher than those of the before–8-week group). The length-of-stay comparisons follow a similar pattern.

We now turn to the results of the regression analysis. The coefficient estimates and *t* statistics for model 1 are shown in Table 3. For a one-tailed test, *t* statistics above 1.65 are statistically significant at $P < .05$. This table reports four sets of estimates: two dependent variables (log of wage and inflation adjusted charges and log of length of stay) for two data sets (all observations and the matched data). Because the dependent variables are in logarithms, for each continuous variable (e.g., age) in the regression model, each coefficient is the estimate of the percentage effect on the dependent variable of a one-unit change in the independent variable. For binary variables, the simple interpretation that the coefficients show the percentage difference in charges or length of stay for the presence or absence of that factor is biased.[27] The actual percentage effect of the coefficient can be accurately approximated by $\{(\exp(c))-1\}$, where c is the coefficient estimate. The size of this bias increases with the coefficient estimate. For a coefficient of 0.25, the true percentage effect is 0.28 (28%); for an estimate of 0.50, the percentage effect is 0.65.

For exposition, we will work through the meaning of the coefficients of the first column of Table 3 (model 1). Because of the adjustments required for binary variables, discussed previously, we report two numbers in the text for each coefficient for all binary variables—the actual estimate and the adjusted estimate, indicated in parentheses. The maternal age effect is very small: hospital charges decreased by 0.6% for each year that maternal age increased. However, the coefficient is not statistically significant; therefore, this effect should be considered to be zero. Compared with White's class B (the reference group), cases in White's class A incurred charges that were 47.0% (37.5%) lower, cases in White's class C incurred charges that averaged 41.5% (51.5%) higher, and cases in White's class D+ incurred charges that averaged 54.5% (72.4%) higher. Charges to MediCal mothers averaged 30.9% (36.2%) higher than those covered by private insurance.

Charges for Black mothers averaged 25.4% (28.9%) more than those for non-Black mothers. A key finding is that mothers and babies in the CDAPP incurred charges that averaged 30.3% (35.3%) lower than charges for the control cases. Hospital charges and length of stay increased rapidly as birth weight decreased from the normal-weight range of 2500 g to 4000 g. However, there were no differences in charges and length of stay between normal-weight cases and cases with birth weights over 4000 g. The adjusted R^2 of 0.259 indicates that the model explains 25.9% of the variance in the logarithms of total charges.

Because the effects of the control variables are stable across models, we report the coefficient estimates for the program effects of the other models in Table 4. To facilitate comparison with models 2 and 3, we repeat the estimates of the program effect for model 1 in Table 4. For model 1, CDAPP cases incurred charges that were 30.3% (35.3%) lower for the entire sample and 24.7% (28.0%) lower for the matched sample. CDAPP cases' length of stay was 26.2% (30.0%) shorter for the entire sample and 19.2% (21.1%) shorter for the matched sample. When the regression results are converted to dollars and days (by applying Duan's correction[26]), they represent average reductions for each mother and baby of $7692 and $6312 in charges, and reductions in length of stay of 5.1 and 4.0 days, for the entire sample and the matched sample, respectively.

Model 2 breaks the program effects into early and late enrollment. Early enrollment reduced length of stay by 39.6% (48.6%) and charges by 34.5% (41.2%) in the entire sample. In the matched sample, early enrollment reduced length of stay by 34.6% (41.4%) and charges by 29.5% (34.4%) not significant (NS). Late enrollment reduced charges by 28.9% (33.5%) and length of stay by 21.2% (23.6%) for the entire sample, and reduced charges by 23.3% (26.2%) and length of stay by 14.0% (15.0%) (NS) for the matched sample. When these results are transformed into dollar or days terms, they represent reductions for early enrollees of $8759 and $7539 in charges and 7.7 days and 7.2 days in length of stay for the entire sample and the matched sample, respectively. For late enrollees, the corresponding reductions are $7344 (entire sample) and $5941 (matched sample) and 3.1 days (entire sample) and 2.9 days (matched sample).

Model 3 was used to examine the differential effects of the program across severity of diabetes. Interpretation of the program effects is more complex for this model, because the effects of White's class must be included. The excluded category is control cases in White's class A or B. To compare a group with the excluded category, the White's class variable is added to the interaction variable between the CDAPP group and White's class. For the most severe cases of diabetes (D+), charges were 51.0% (66.6%) and 65.8% (93.2%) lower than charges for control cases for the matched sample and the entire sample, respectively. Length of stay was 29.4% (34.2%) (NS) and 48.8% (62.9%) shorter for the matched sample and the entire sample, respectively. For class D+, the retransformed charges for the CDAPP group were $16 766 lower than charges for the control group. Except for length of stay for the matched sample for White's classes A and B, the coefficient estimates for the less severe diabetics are not statistically significant. In sum, the savings were larger for the most severe cases of diabetes (those in class D+). However, these numbers must be treated with some caution owing to the reduced sample sizes, especially for class D+.

These models were reestimated with the maternal and infant components of charges and length of stay as separate dependent variables. Overall, the program had similar effects for mothers and infants. We also found no difference in the estimated program effects between Blacks and non-Blacks and between MediCal and other payors.

TABLE 4—Estimated Coefficients for Program Effects, Regression Models 1, 2, and 3

	Cases		Matched Cases	
	Charges	Length of Stay	Charges	Length of Stay
		Model 1		
CDAPP enrollment	−0.3026	−0.2618	−0.2469	−0.1916
	(−2.821)**	(−2.827)**	(−2.068)*	(−1.771)*
		Model 2		
CDAPP early enrollment	−0.3448	−0.3960	−0.2953	−0.3462
	(−1.919)*	(−2.636)**	(−1.504)	(−1.953)*
CDAPP late enrollment	−0.2891	−0.2116	−0.2327	−0.1401
	(−2.475)**	(−2.063)*	(−1.817)*	(−1.170)
		Model 3		
Interaction classes A & B	−0.0871	−0.1592	−0.2298	−0.2871
	(−0.627)	(−1.305)	(−1.514)	(−2.062)*
Interaction class C	−0.2191	−0.0778	0.0667	0.2510
	(−0.922)	(−0.390)	(0.220)	(0.939)
Interaction class D+	−0.6584	−0.4877	−0.5102	−0.2938
	(−2.736)**	(−2.321)**	(−2.033)*	(−1.192)

Note. These models control for mother's age, White's class of diabetes, insurance (MediCal), race (Black), and birth weight. The estimated dependent variables are log charges and log length of stay for mothers and babies. *T* statistics are given in parentheses.
*$P \leq .05$ (one-sided).
**$P \leq .01$ (one-sided).

The savings estimated from the regression models were discounted to make them directly comparable with program costs. Table 5 shows these discounted savings. Turning first to the results for model 1, we report the most conservative estimate (the matched group) of the savings generated by the CDAPP. For the Pre/Post CDAPP group, the savings were $6072 per participant. If all women went into the program before conception (preconception) the savings would be $5624 per participant, because the CDAPP costs are higher for this group. Model 2, which shows estimated program savings for early and late enrollees, indicates that savings for early enrollees (those who enrolled within 8 weeks of conception) were $7253, compared with $5715 for late enrollees. The results for the entire sample are similar and are slightly larger in most cases. The larger program costs under the All Pre assumption lead to slightly smaller net savings. Model 3 shows that the savings increased with the severity of diabetes. The counterintuitive result for White's class C in the matched sample may be due to the relatively small sample size (NS).

The cost-effectiveness of the CDAPP is calculated by comparing the discounted savings shown in Table 5 with the program costs. This is done by subtracting the program costs from the discounted savings to obtain net savings. To make this calculation meaningful, we allocated to the CDAPP all expenditures incurred during

TABLE 5—Discounted Savings (in dollars) for Models 1, 2, and 3

	All Cases		Matched Cases	
	All Pre	Pre/Post	All Pre	Pre/Post
Model 1	6854	7400	5624	6072
Model 2				
Early	7804	8426	6717	7253
Late	6544	7065	5293	5715
Model 3				
Classes A & B	1976	2134	5201	5615
Class C	4971	5367	(−1509)	(−1630)
Class D+	14 939	16 129	11 547	12 468

Note. An 8% discount rate was used. All pre-conception enrollments were discounted at 0.891. The mixture of pre and post enrollments was discounted at 0.962.

TABLE 6—Cost-effectiveness Ratios[a] for Models 1, 2, and 3

	All Cases		Matched Cases	
	All Pre	Pre/Post	All Pre	Pre/Post
Model 1	3.69	6.32	3.03	5.19
Model 2				
Early	4.20	7.20	3.62	6.19
Late	3.52	6.03	2.85	4.88
Model 3				
Classes				
A & B	1.06	1.82	2.80	4.80
Class C	2.68	4.58	−0.81	−1.39
Class D+	8.04	13.77	6.22	10.65

Note. Assumes 70% of program participants ultimately give birth to live infants. All pre: $1300 to $1857 per CDAPP participant. Pre/post: $820 to $1171 per CDAPP participant.

[a] Total Discounted Savings/Program Costs (these are costs of program)

our study period. This procedure produces a downward bias in our savings (i.e., savings will be lower), because additional savings will come as other CDAPP participants have additional children. Both the Pre/Post cost of $1171 and the All Pre cost of $1857 are markedly less than the estimated savings for all of the models.

Alternatively, the gross savings can be divided by the program costs to obtain a cost-effectiveness ratio. This ratio measures the short-term savings per dollar spent on the CDAPP. These ratios are presented in Table 6. We find that for the matched sample, the CDAPP has a cost-effectiveness ratio of 5.19; this means that for every $1 spent on the CDAPP there is an estimated savings of $5.19. These ratios also show the marginal effectiveness of applying the program to different subsets of the study population. The largest returns will be realized if the program targets early enrollees and more severe cases of diabetes. The returns are smaller for late enrollees, and smaller and not statistically significant for less severe cases of diabetes.

Discussion

The CDAPP has a statistically significant effect that lowers hospital charges and length of stay for diabetic mothers and their infants. Moreover, the program is cost-effective and returns over $5 for every $1 spent. The conservative assumptions used in the analysis suggest that the actual savings due to the CDAPP are larger than our estimates.

Comparisons of demographic characteristics of the women who gave birth in the CDAPP (from preliminary program status reports) and women who gave birth in the general population in California in 1986[20] show that these groups are quite similar. Data on three pregnancy risk factors—maternal age, prenatal care in first trimester, and race—were considered. In the high-risk age group, 4.0% of all California births were to women under 18 years old, compared with 3.6% of births in the CDAPP; 8.8% of all births were to women over 35 years old, compared with 21.4% of births in the CDAPP. Of all births, 74% were to women who began medical care in the first trimester of pregnancy; 67.2% of births in the CDAPP were to women with overt diabetes who joined the CDAPP in the first trimester. Of all births, 10% were to Black women; 7% of births among the overt diabetics in the CDAPP were to Black women. These comparisons suggest that our results are quite generalizable to California mothers.

The program is more cost-effective if the mothers enroll before conception, or shortly thereafter. Our estimates show that early enrollees incur lower hospital charges and have shorter stays. Although late enrollees also incur lower charges and have shorter stays, the CDAPP effects are smaller. It is conceivable that when the CDAPP is fully implemented, early program participation will be increased. If this occurs, our estimates suggest that the program savings will increase. The program savings for early enrollees are 10% to 15% larger than those for late enrollees. Declines in length of hospital stay are even larger. Further, the program is more cost-effective for those mothers with more severe diabetes (White's class D+).

The lack of statistical significance in model 3 for White's classes A and B and class C does not mean that the program is not effective for these cases. Given the relatively large variation in costs and length of stay and the smaller expected effects for less severe diabetics, it is possible that our sample was not of sufficient size to enable us to detect differences for these groups. Further study is needed to determine whether the program actually reduces costs and length of stay for these less severe cases.

Our overall finding is that the CDAPP yields substantial short-term economic savings. If 2598 babies are born to women with overt diabetes in California in 1990, as estimated by the California Department of Health Services in 1987, the estimated yearly savings attributable to the CDAPP will be between $14 million and $19 million, depending on how many women enroll in the program before conception. In addition to the short-term economic savings of the program, there are unmeasured savings that have not been captured in our estimates. Improved glucose management will likely result in improvements in the future health status of the mothers, assuming that they retain the knowledge they gained while in the program. Reductions in the severity of birth complications will likely result in fewer anomalies and other handicaps, yielding long-term savings due to reduced medical and educational expenses and increased future productivity. With a more normal pregnancy and a healthy baby, a working mother can remain in the work force longer and return to work sooner. The reduction in perinatal problems may minimize disruption of the father's ability to work, as well. In addition, there are many nonmonetary benefits to healthy pregnancy outcomes. □

Acknowledgments

This study was funded by the California Department of Health Services, Maternal and Child Health Branch (contract 88-93733).

The support, time, and energy of many individuals contributed to the success of this study. We would like to express our gratitude to Diana Lee, RD, MPH, Joe Brazie, MD, MPH, Donna Dahrouge, MPH, and Jan Pierce, who all played essential roles in the initiation and completion of the study. We also wish to thank Larry M. Cousins, MD, Jack Kitzmiller, MD, and Carol Korenbrot, PhD, for their useful comments and suggestions on an earlier draft. Lastly, we extend our gratitude to the staff of the medical and billing records departments of all of the hospitals we visited for providing the data needed for this study.

References

1. Russel LB. *Evaluating preventive care.* Washington, DC: The Brookings Institute; 1987.
2. Scheffler RM, Paringer L. A review of the economic evidence on prevention. *Med Care.* 1980;18:473–484.
3. Miller E, Hare JW, Cloherty JP. Elevated maternal hemoglobin A1c in early pregnancy and major congenital anomalies in infants of diabetic mothers. *N Engl J Med.* 1981;304:1331–1334.
4. Cousins L. Congenital anomalies among infants of diabetic mothers. *Am J Obstet Gynecol.* 1983;147:333–338.
5. Lowy C, Beard RW, Goldschmidt J. Congenital malformations in babies of diabetic mothers. *Diabetic Med.* March 1986:458–462.
6. Steele JM, Johnstone FD, Hepurn DA, Smith AF. Can pre-pregnancy care of diabetic women reduce the risk of abnormal babies? *Br Med J.* 1990;301:1070–1074.
7. Kitzmiller JL, Gavin LA, Gin GD, Jovanovic-Peterson L, Main EK, Zigrang WD. Preconception care of diabetes: glycemic control prevents congenital anomalies. *JAMA.* 1991;265:731–736.
8. Centers for Disease Control. Public health guidelines for enhancing diabetes control through maternal and child health programs. *MMWR.* 1986;35:201–213.
9. Hollingsworth DR, Jones OW, Resnik R. Expanded care in obstetrics for the 1980's: preconception and early conception counseling. *Am J Obstet Gynecol.* 1984;149:811.
10. Institute of Medicine, Division of Health Promotion and Disease Prevention, Committee to Study the Prevention of Low Birth Weight, *Preventing Low Birth Weight.* Washington, D.C. National Academy Press; 1985.
11. Korenbrot CC. Risk Reduction in Pregnancies of Low Income Women: Comprehensive Prenatal Care through the OB Access Project. Mobius. 1984;4:34–43.
12. Klebe JG, Espersen T, Allen J. Diabetes mellitus and pregnancy. *Acta Obstet Gynecol Scand.* 1986;65:235–240.
13. Fuhrman K, Reiher H, Semmler K. Prevention of congenital malformations in infants of insulin-dependent diabetic mothers. *Diabetes Care.* 1983;6:219–223.
14. Mills JL, Baker L, Goldman AS. Malformations in infants of diabetic mothers occur before the seventh gestational week. *Diabetes.* 1979;25:292–293.
15. Freinkel N, Dooley SL, Metzger BE. Care of the pregnant woman with insulin-dependent diabetes mellitus. *N Engl J Med.* 1985;11:96–101.
16. Landon MB, Gabbe SG, Piana R, Mennuti MT, Main EK. Neonatal morbidity in pregnancy complicated by diabetes mellitus: predictive value of maternal glycemic profiles. *Am J Obstet Gynecol.* 1987;156:1089–1095.
17. Mills JL, Knopp RH, Simpson JL, et al. Lack of relation of increased malformation rates in infants of diabetic mothers to glycemic control during organogenesis. *N Engl J Med.* 1988;318:671–675.
18. Key TC, Giuffrida R, Moore TR. Predictive value of early pregnancy glycohemoglobin in the insulin-treated diabetic patient. *Am J Obstet Gynecol.* 1987;156:1096–1099.
19. Goldman JA, Dicker D, Feldberg D, Yeshaya A, Samuel N, Karp M. Pregnancy outcome in patients with insulin-dependent diabetes mellitus with preconceptional diabetes control: a comparative study. *Am J Obstet Gynecol.* 1986;155:293–297.
20. Williams RL, Rust FP. Maternal and Child Health Data Base 1982–1986: Descriptive, Narrative and Statistical Appendix. Report for the Maternal and Child Health Branch, State of California. Department of Health Services, MCH Contract #90-10485.
21. US Dept of Health and Human Services, National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes.* 1979;28:1039–1057.
22. *Guidelines for Care.* California Diabetes and Pregnancy Program. Sacramento, Calif: California Department of Health Services, Maternal and Child Health Branch; 1987.
23. Lin CK, Kuo PL, Liu HC. Clinical analysis of infants of diabetic mothers. *Acta Paediatric Sin.* 1989;30:233–239.
24. Berk MA, Miimouni F, Miodovnik M, Hertzberg V, Valuck J. Macrosomnia in infants of insulin-dependent diabetic mothers. *Pediatrics.* 1989;83:1029–1034.
25. Phibbs CS, Phibbs RH, Pomerance JJ, Williams RL. An alternative to diagnostic related groups for newborn intensive care. *Pediatrics.* 1986;78:829–836.
26. Duan N. Smearing estimate: a nonparametric retransformation method. *J Am Stat Assoc.* 1983;78:605–610.
27. Halvorsen R, Palmquist R. The interpretation of dummy variables in semilogarithmic equations. *Am Econ Rev.* 1980;70:474–475.

Maternal Smoking and the Risk of Polyhydramnios

Wendie Myhra, MSN, MPH, Margaret Davis, DVM, MPH, Beth A. Mueller, DrPH, and Durlin Hickok, MD, MPH

ABSTRACT

Background. Washington State birth certificates were used to conduct a population-based case-control study to assess the possible association of maternal smoking with polyhydramnios.

Methods. All singleton births complicated by polyhydramnios (n = 557) were identified from the vital records for the years 1984 to 1987. For comparison, 1671 records were randomly selected for the same years from singleton births uncomplicated by polyhydramnios.

Results. Women who reportedly smoked prenatally were found to be at increased risk for polyhydramnios (relative risk [RR] = 1.7, 95% confidence interval [CI] = 1.5–2.1, adjusted for marital status, maternal age, and parity). When women with conditions known to be associated with polyhydramnios were excluded, the risk for those who smoked prenatally remained elevated (RR = 1.8, 95% CI = 1.1–2.3).

Conclusion. Overdistention of the uterus from polyhydramnios may cause a variety of pregnancy complications. The observed association of smoking with polyhydramnios may be a further indication for public health interventions aimed at preventing smoking during pregnancy. (*Am J Public Health.* 1992;82:176–179)

Introduction

Polyhydramnios, an excess of amniotic fluid during pregnancy, is a condition whose causes and implications are not completely understood. Although it is known that polyhydramnios occurs more often in women with established and gestational diabetes, multiple pregnancies, and pregnancy complications such as preeclampsia, as well as in women whose babies have congenital malformations, the mechanism by which it occurs is poorly understood. In addition, polyhydramnios is reported with greater frequency than can be explained by these conditions alone.[1,2] Thus, to understand better the possible causes of polyhydramnios, we carried out a population-based case-control study using birth certificate data.

Methods

A total of 277 043 singleton births were recorded on the Washington State birth certificate tapes from 1984 through 1987. Of these, 557 pregnancies were reported to be complicated by polyhydramnios. A control group consisting of 1671 singleton births was randomly selected from among those births for which polyhydramnios was not indicated as a complication. The Washington State birth certificate contains a check-box format, which has been shown to provide more complete reporting of birth complications.[3] This information was collected, as were data on maternal and infant characteristics, complications of pregnancy and labor, and congenital malformations.

A stratified analysis was conducted using the Mantel-Haenzel estimator of the odds ratio, and 95% confidence intervals (CIs) were calculated using Cornfield's method.[4] Several variables were considered as potential confounders; these included maternal age (under 35 years, 35 years and over), race (White, Black, other), marital status, smoking status during present pregnancy, parity (zero, one or more), urban or rural residence, income level of census tract of residence, trimester of first prenatal care, sex of infant, complications of pregnancy, congenital malformations, prior fetal deaths (fewer than 20 weeks gestation, 20 to 36 weeks gestation, more than 36 weeks gestation), and season of birth. However, only those variables that affect the relative risk (RR) estimate were retained for analysis.

The initial analysis was carried out using the entire study population. The majority of subsequent analyses were restricted to White women only because there were very few non-White subjects and the small number made attempts to control for race unsuccessful.

Results

A greater proportion of cases than of controls were aged 35 years or older, un-

The authors are all with the Department of Epidemiology at the University of Washington in Seattle. Additionally, Beth Mueller is in the Public Health Sciences Division of the Fred Hutchinson Cancer Research Center, and Durlin Hickok is with the Department of Perinatology at the Swedish Hospital Medical Center in Seattle.

Requests for reprints should be sent to Dr. Beth Mueller, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, 1124 Columbia Street MP-381, Seattle, WA 98104.

This paper was submitted to the journal May 2, 1990, and accepted with revisions September 20, 1991.

married, and of higher parity (Table 1). The proportion of cases with gestational or established diabetes or with preeclampsia was also higher than that for controls, and a greater proportion of cases than controls experienced complications such as abruptio placentae or placenta previa, or delivered a baby with a congenital anomaly. Although the great majority of both cases and controls were White, the proportion of Black cases was nearly twice that of Black controls, and among Hispanics there was a relatively lower proportion of cases (1.6%) than of controls (2.6%).

Among Whites, the proportion of cases who smoked (34.5%) was greater than that for controls (23.8%), resulting in an increased risk among cases for polyhydramnios (RR = 1.7, 95% CI = 1.5–2.1, adjusted for maternal age, marital status, and parity) (Table 2). White women who were 35 years of age or older were at increased risk for polyhydramnios (RR = 2.6, 95% CI = 1.6–3.2, adjusted for marital status, smoker status, and parity), relative to those aged 20 to 34 years of age. The risk for polyhydramnios also increased with increasing parity. White women having one previous birth compared with those having no previous births were at a slightly increased risk for polyhydramnios (RR = 1.2, 95% CI = 1.0–1.5), and those with three or more previous births had nearly twice the risk for polyhydramnios (RR = 1.9, 95% CI = 1.2–3.3). (These risks were adjusted for maternal age, smoker status, and marital status.)

After excluding women with preeclampsia, established and gestational diabetes, and babies with congenital malformations, the risk of polyhydramnios associated with maternal smoking among White women was 1.8 (95% CI = 1.1–2.3) (Table 3). The risks associated with maternal age of 35 years or older increased slightly as well, to 3.1 (95% CI = 1.9–5.3). The risks associated with higher parity remained consistently elevated.

Women with established diabetes were at a greatly increased risk for polyhydramnios (RR = 23.9, 95% CI = 5.2–957.0). Those with gestational diabetes had a risk of 7.3 (95% CI = 3.2–18.3), and those with preeclampsia had a risk of 2.8 (95% CI = 1.8–4.9). (All risk estimates were adjusted for maternal age and the presence of fetal malformations.) The risk estimates for gestational diabetes and preeclampsia were also adjusted for smoker status.

In contrast to the risk estimate for polyhydramnios associated with smoking among White women, the risk among Black women was less than 1, even after adjusting for maternal age (RR = 0.6, 95% CI = 0.1–1.9). Adjusting for parity, marital status, gravidity, diabetes, preeclampsia, or the presence of congenital malformations did not effect the risk estimate. The number of Black cases and controls was too small to adjust for all variables simultaneously.

TABLE 1—Maternal, Infant, and Pregnancy Characteristics of Polyhydramnios Cases and Controls[a]

	Cases (n = 557)		Controls (n = 1671)	
	n	%	n	%
Age in years				
<20	58	10.4	178	10.7
20–34	428	76.8	1393	83.4
≥35	71	12.8	99	5.9
Race				
White	474	85.7	1453	87.6
Black	35	6.3	54	3.3
Other	44	8.0	151	9.1
Unmarried	120	21.6	299	17.9
Smoked during pregnancy	166	33.3	368	23.6
Parity				
No previous	189	36.0	699	44.0
One previous	164	31.2	541	34.0
Two previous	103	19.6	218	13.7
Three plus previous	69	13.2	132	8.3
Pregnancy complications				
Est. Diabetes	15	2.7	3	0.2
Gest. Diabetes	37	6.6	13	0.8
Preeclampsia	46	8.3	55	3.3
Congenital Anomaly	102	18.3	27	1.6
Abruptio Placentae	8	1.4	13	0.8
Placenta Previa	5	0.9	3	0.2

[a]Data are missing on age for 1 control, race for 4 cases and 13 controls, marital status for 2 cases and 4 controls, smoker status during pregnancy for 58 cases and 112 controls, and parity for 32 cases and 81 controls.

TABLE 2—Risks Associated with Maternal Characteristics of White Polyhydramnios Cases and Controls[a]

	Cases (n = 420)		Controls (n = 1343)			
	n	%	n	%	RR	95% CI
Age in years						
<20	48	11.4	131	9.8	1.3	(0.8–2.3)
20–34	318	75.7	1138	84.7	1.0	(–)
≥35	54	12.9	74	5.5	2.6	(1.6–3.2)
Unmarried	83	19.8	209	15.6	1.3	(0.8–1.6)
Smoked during pregnancy	145	34.5	320	23.8	1.7	(1.5–2.1)
Parity[b]						
No previous	146	34.8	600	44.7	1.0	(–)
One previous	137	32.6	462	34.4	1.2	(1.0–1.5)
Two previous	83	19.8	179	13.3	1.8	(1.1–2.3)
Three plus previous	54	12.8	102	7.6	1.9	(1.2–3.3)

Note. RR = relative risk; CI = confidence interval.
[a]All risk estimates are adjusted for all other variables in table.
[b]Test for trend, $\chi^2 = 17.28$, $P < .001$.

Discussion

These results suggest that maternal smoking during pregnancy is associated with an increased risk for polyhydramnios, even in the absence of other recognized risk factors. To our knowledge, this association has not been reported previously. Our finding of an increased risk for polyhydramnios associated with the pres-

TABLE 3—Risks Associated with Maternal Characteristics of White Polyhydramnios Cases and Controls Excluding All Pregnancies with Preeclampsia, Established Diabetes, Gestational Diabetes, or Congenital Malformations[a]

	Cases (n = 286)		Controls (n = 1271)			
	n	%	n	%	RR	95% CI
Age in years						
<20	33	11.5	121	9.5	1.4	(0.9–2.5)
20–34	213	74.5	1084	85.3	1.0	(–)
≥35	40	14.0	66	5.2	3.1	(1.9–5.3)
Unmarried	60	21.0	198	15.6	1.4	(0.9–1.7)
Smoked during pregnancy	107	37.4	308	24.2	1.8	(1.1–2.3)
Parity[b]						
No previous	91	31.8	553	43.5	1.0	(–)
One previous	98	34.3	447	35.2	1.3	(0.8–1.6)
Two previous	57	19.9	172	13.5	1.9	(1.2–3.3)
Three plus previous	40	14.0	99	7.8	2.1	(1.3–3.7)

Note. RR = relative risk; CI = confidence interval.
[a]All risk estimates are adjusted for all other variables in table.
[b]Test for trend, $\chi^2 = 16.54$, $P < .001$.

ence of gestational diabetes, established diabetes, preeclampsia, and congenital malformations is consistent with earlier reports.[1,2,5–7]

In addition, these data suggest that the risk of polyhydramnios increases with increasing parity. Women with one previous birth are at a slightly increased risk for polyhydramnios, and those with three or more previous births are at nearly twice the risk, compared with those with no prior births; this remains true even after adjusting for maternal age, smoking status, and marital status. It is possible that this apparent association occurs because women with prior births may have had polyhydramnios as a complication of a previous pregnancy. Whether a history of polyhydramnios increases the likelihood of developing polyhydramnios in subsequent pregnancies is unknown. However, women with histories of gestational diabetes, of carrying a malformed fetus, and of other complications in pregnancy are more likely to experience these complications in subsequent pregnancies than are women with no such histories. Thus, it is plausible that the association with parity seen in this study may, in part, be due to an association with these adverse occurrences in prior pregnancies.

Increased maternal age was also a risk factor for polyhydramnios, even after adjusting for parity, smoking, and marital status and excluding pregnancies with complications associated with polyhydramnios and those resulting in the birth of a baby with congenital malformations.

Several biological mechanisms might explain the apparent association between cigarette smoking and polyhydramnios. There is experimental evidence that lactogenic hormones are involved in regulating fluid flux across placental membranes.[8] A decrease in human placental lactogen has been demonstrated in smokers,[9] and chronic idiopathic polyhydramnios has been shown to be associated with a deficit in chorionic laeve prolactin receptors.[10] A deficit in human placental lactogen could have the same effect as a deficit in receptors.

The relationship between smoking and polyhydramnios may also be mediated by the hormone angiotensin, which functions in the control of vascular volume. Angiotensin I (AI) is normally converted to angiotensin II (AII) in the lung by angiotensin-converting enzyme (ACE). AII is associated with peripheral vasoconstriction. Human and animal studies have shown that smoking elevates ACE levels and enhances activation of the angiotensin system.[11–13] A connection between the angiotensin system and the development of polyhydramnios is supported by animal models in which infusion of AI into fetal lambs induced polyhydramnios.[14]

Finally, polyhydramnios may be linked to changes that occur in placental vasculature as a result of cigarette smoking. Nicotine reduces placental blood flow in humans and animals.[15–17] Microscopic placental examination has demonstrated atrophic and hypovascular changes in the placental villae of smokers, suggesting that prenatal smoking results in impaired uteroplacental circulation.[9]

The relatively "soft" nature of the method by which polyhydramnios is diagnosed constitutes a limitation to this study. *Williams Obstetrics*[18] defines polyhydramnios as more than 2000 mL of amniotic fluid. Until very recently,[19] no attempts were made to measure normal amniotic fluid volumes in a large population of pregnant women.[7,19] Although available,[20–22] objective sonographic criteria for the diagnosis of polyhydramnios may not always be applied. Because of this, the diagnosis of polyhydramnios often relies on the subjective clinical judgment of each woman's health care provider. Our method of case ascertainment is limited to a yes/no variable from the birth certificate, without any knowledge of how the diagnosis was made. Births with labor complications may be more likely than others to be recorded as having been complicated by polyhydramnios. Ultrasound examination, often indicated for conditions of pregnancy known to be associated with polyhydramnios, is also a method of case finding. Both of these mechanisms might possibly result in underreporting of idiopathic polyhydramnios. However, this would have led to a conservative bias in our risk estimate.

Our indicator of a woman's smoker status is also limited in that it does not allow any measurement of the dose of cigarettes smoked or of the duration of smoking during pregnancy. It is most likely that our variable for smoking reflects a woman's smoker status at the time of her first prenatal visit, as this is when this item of her medical record is checked. Many smokers change their behavior while pregnant,[23] so it is likely that some women who stopped smoking during their pregnancies were misclassified in the smoker category. Again, this would have had the effect of biasing our risk estimates toward the null.

Another limitation is that our data were limited to live births. Fetal losses and elective terminations of pregnancy were not included in the study, resulting in possible bias. Spontaneous abortion is often associated with fetal anomalies, and it shortens pregnancies that, had they continued, might have been complicated by polyhydramnios. Induced abortions following early detection of congenital malformations by alpha-fetoprotein screening or amniocentesis may also result in a decreased number of polyhydramnios cases with malformations. These two mechanisms would result in an underestimation of the true relative risk associated with the presence of malformations.

In this population, the prevalence of polyhydramnios at birth was 0.2% com-

pared with ranges in other reports from 0.2% to 1.7%. The observed variance in prevalence may be due to the subjective nature of the diagnosis and to the differences between populations studied.[1,5,6,24] Most prior investigations of polyhydramnios have been limited to hospital or clinic-based series, and our population-based figure appears to lie well within this range. Further population-based studies, using strict criteria, would provide a better understanding of potential risk factors and outcomes associated with polyhydramnios.

Overdistention of the uterus from polyhydramnios may cause a variety of pregnancy complications, including inefficient uterine action resulting in preterm labor, abruptio placenta, abnormal fetal lie, and dystocia.[6,25] Therefore, polyhydramnios may elevate a woman's risk for complications even in the absence of other known risk factors.[4,23] The possible association of smoking and polyhydramnios is an added indication for public health interventions aimed at preventing smoking during pregnancy. □

References

1. Cardwell MS. Polyhydramnios: a review. *Obstet Gynecol Surv.* 1987;42:612–617.
2. Hill LM, Breckle R, Thomas ML, Fries JK. Polyhydramnios: ultrasonically detected prevalence and neonatal outcome. *Obstet Gynecol.* 1987;69:21–25.
3. Frost F, Starzyk P. Birth complication reporting: the effect of birth certificate design. *Am J Public Health.* 1984;74:505–506.
4. Schlesselman JJ. *Case Control Studies: Design, Conduct, Analysis.* New York, NY: Oxford University Press; 1982:177–192.
5. Varma TR, Bateman S, Patel RH, Chamberlain GVP, Pillai U. The relationship of increased amniotic fluid volume to perinatal outcome. *Int J Gynecol Obstet.* 1988;27:327–333.
6. Boylan P, Parisi V. An overview of polyhydramnios. *Semin Perinatol.* 1986;10:136–141.
7. Leucht W, Rabe D, Hendrik HJ, Grischke EM, Schmidt W. Sonographic evaluation of the amount of amniotic fluid: polyhydramnios—significance for the course of pregnancy and labor. *Geburtshilfe Frauenheilkd.* 1986;46:157–161.
8. Leontic EA, Tyson JE. Prolactin and fetal osmoregulation: water transport across isolated human amnion. *Am J Physiol.* 1977;232:R124–R127.
9. Mochizuki M, Maruo T, Mosuki K, Ohtsu T. Effects of smoking on fetoplacental-maternal system during pregnancy. *Am J Obstet Gynecol.* 1984;149:413–420.
10. Healy DL, Herington AC, O'Herlihy C. Chronic idiopathic polyhydramnios: evidence for a defect in the chorionic laeve receptor for lactogenic hormones. *J Clin Endocrinol Metab.* 1983;56:520–523.
11. Laustiola KE, Lassila R, Nurmi AK. Enhanced activation of the renin-angiotensin-aldosterone system in chronic cigarette smokers: a study of monozygotic twin pairs discordant for smoking. *Clin Pharmacol Ther.* 1988;44:426–430.
12. Sugiyama Y, Yotsumoto H, Okabe T, Takaku F. Measurement of angiotensin-converting enzyme activity in intact human alveolar macrophages and effect of smoking. *Respiration.* 1988;53:153–157.
13. Sugiyama Y, Yotsumoto H, Takaku F. Increase of serum angiotensin-converting enzyme level after exposure to cigarette smoke and nicotine infusion in dogs. *Respiration.* 1986;49:292–295.
14. Anderson DF, Faber JJ. Animal model for polyhydramnios. *Am J Obstet Gynecol.* 1989;160:389–390.
15. Morrow RJ, Ritchie JWK, Bull SB. Maternal cigarette smoking: the effects on umbilical and uterine blood flow velocity. *Am J Obstet Gynecol.* 1988;159:1069–1071.
16. Suzuki K, Horiguchi T, Comas-Urrutia AC, Mueller-Heubach E, Morishima HO, Adamsons K. Pharmacologic effects of nicotine upon the fetus and mother in the rhesus monkey. *Am J Obstet Gynecol.* 1971;111:1092–1101.
17. Lehtovirta P, Forss M. The acute effect of smoking on intervillous blood flow of the placenta. *Br J Obstet Gynaecol.* 1978;85:729–731.
18. Pritchard JA, MacDonald PC, Gant NF, eds. *Williams Obstetrics.* Norwalk, Conn: Appleton-Century-Crofts; 1985.
19. Brace RA, Wolf EJ. Normal amniotic fluid volume changes throughout pregnancy. *Am J Obstet Gynecol.* 1989;161:382–388.
20. Crowly P, O'Herlihy C, Boylan P. The value of ultrasound measurement of amniotic fluid in the management of prolonged pregnancies. *Br J Obstet Gynaecol.* 1984;91:444–448.
21. Hoddick WK, Callen PW, Filly RA. Ultrasonographic determination of qualitative amniotic fluid volume in intrauterine growth retardation: reassessment of the 1 cm rule. *Am J Obstet Gynecol.* 1984;149:758–762.
22. Rutherford SE, Phelan JP, Smith CV, Jacobs N. The four-quadrant assessment of amniotic fluid volume: an adjunct to antepartum fetal heart rate testing. *Obstet Gynecol.* 1987;70 (part 1):353–356.
23. Fried PA, Watkinson B, Grant A, Knights RM. Changing patterns of soft drug use prior to and during pregnancy: a prospective study. *Drug Alcohol Depend.* 1980;6:323–343.
24. Hill LM. Resolving polyhydramnios: a sign of improved fetal status. *Am J Perinatol.* 1988;5:61–63.
25. Zamah NM, Gillieson MS, Walters JH, Hall PF. Sonographic detection of polyhydramnios: a five-year experience. *Am J Obstet Gynecol.* 1982;143:523.

A Cost-effectiveness Analysis of Three Staffing Models for the Delivery of Low-Risk Prenatal Care

Elaine A. Graveley, DBA, RN, and John H. Littlefield, PhD

ABSTRACT

Background. Health care costs are increasing at more than twice the rate of inflation, thus, public officials are seeking safe and economic methods to deliver quality prenatal care to poor pregnant women. This study was undertaken to determine the relationship between the cost and effectiveness of three prenatal clinic staffing models: physician based, mixed staffing, and clinical nurse specialist with physicians available for consultation.

Methods. Maternal and neonatal physiological outcome data were obtained from the hospital clinical records of 156 women attending these clinics. The women were then interviewed concerning their satisfaction with their prenatal care clinic. The financial officer from each clinic provided data on the clinic staffing costs and hours of service.

Results. There were no differences in outcomes for the maternal–neonatal physiological variables, although newborn admission to the Neonatal Intensive Care Unit (NICU) approached significance among the clinics. The clinic staffed by clinical nurse specialists had the greatest client satisfaction and the lowest cost per visit.

Conclusions. The use of clinical nurse specialists might substantially reduce the cost of providing prenatal care while maintaining quality, and might thereby save valuable resources. (*Am J Public Health.* 1992;82:180–184)

Introduction

With health care costs growing at more than twice the rate of inflation, public officials are seeking safe and economic methods to deliver quality health care to indigent populations. The percentage of indigent individuals who receive quality health care is decreasing. Lewin[1] reported that, from 1976 to 1983, Medicaid covered a decreasing number of the poor (of the population with incomes below the poverty line, 65% were covered by Medicaid in 1976 and 53% in 1983) and that poor pregnant women have been increasingly underserved as a result of budget cuts and policy changes.

The relationship between low-birthweight (LBW) babies and the absence of prenatal care is well documented.[2–4] The Office of Technology Assessment (OTA) estimated that for every LBW birth averted, between $14 000 and $30 525 in newborn and long-term health care costs could be saved.[5] OTA also analyzed the cost-effectiveness of expanding Medicaid coverage to all poverty-level pregnant women in an effort to provide them with early prenatal care. Their conclusion was that such expansion would be cost effective and would be a good investment for the nation.

The National Commission to Prevent Infant Mortality (NCPIM) also studied the problem of LBW and proposed universal access to early maternity care for all expectant mothers.[6] However, neither NCPIM nor OTA recommended the type of health care provider for these additional 194 000 women.[5] With our increasingly litigious society, the number of obstetricians willing to serve Medicaid patients is declining. There are, however, other models of prenatal care that need to be studied for their cost-effectiveness: individual or group physician-based clinics and, depending on state laws, nurse-based clinics that tie medical practice with a nurse midwife, nurse practitioner, or clinical nurse specialist. The focus of this study was a comparative cost-effectiveness analysis of three low-risk prenatal clinic staffing models: physician based, mixed staffing (physician, nurse practitioner, registered nurses, nurse aides), and nurse based. Low-risk prenatal care was defined in this study as care given to a pregnant woman "whose complaint or problem has been resolved, or does not pose a threat to life."[7] The results of this study should be helpful to public policymakers who are implementing OTA's and NCPIM's proposal to increase the availability of prenatal care for indigent patients.

Methods

Clinics

The three clinics served predominantly Hispanic clients of low socioeconomic status and referred all women with complications to the Complicated Obstetrical Clinic at the community health center. The women primarily delivered at the hospital operated by the county hospital district.

Clinic 1. This private-not-for-profit entity was funded under Titles 329 and 330 of the Public Service Health Act and served as the physician-based clinic

The authors are with the University of Texas Health Science Center, San Antonio.

Requests for reprints should be sent to John H. Littlefield, PhD, Director, Instructional Development, University of Texas Health Science Center, 7703 Floyd Curl Drive, San Antonio, TX 78284-7948.

This paper was submitted to the journal November 2, 1990, and accepted with revisions July 17, 1991.

(MDC) in this study. Women saw the same physician for all their prenatal care and counseling and received a standard clinic appointment. Fees were charged on a sliding scale, according to the woman's ability to pay.

Clinic 2. This clinic, one of 19 operated by the city health department, served as the mixed-staffing clinic (MSC). The clinic was located within the same catchment area as the MDC. Each patient saw a nurse aide, a registered nurse for prenatal teaching, and the nurse practitioner and/or contract physician at each visit. The clinic operated twice a week from 8 AM to 12 noon. All women received an 8 AM appointment, and prenatal care and counseling were provided for $50 unless the woman was unable to pay.

Clinic 3. This clinic was jointly operated by the county hospital district and the local University Health Science Center's School of Nursing and Department of Obstetrics and Gynecology and served as the nurse-based clinic (RNC). The clinic was located within the county community health center. Three clinical nurse specialists delivered total prenatal care, including referrals to community agencies. Fees were charged on a sliding scale, according to the woman's ability to pay, and the women received a standard appointment time.

Subjects

The sample size, 156 subjects, was determined by a power analysis of 3 months of birth-weight data from women who attended the clinics and delivered at the county hospital ($\alpha = .05$, effect size = .25, power = .8; 3 groups). The study sample was drawn over a 3-month period in 1989 from women who had delivered at the county hospital. To be included, subjects had to: (1) be 18 years of age or older, or emancipated minors; (2) have obtained their prenatal care at one of the three clinics with a minimum of three prenatal visits (three visits was considered the minimum for the woman to have sufficient experience to validly report her satisfaction with her prenatal care clinic); and (3) have delivered within 48 hours of the interview.

Potential subjects were identified from their hospital record and asked to participate in a study of prenatal care clinics. No potential subject was excluded because of language, a Spanish interpreter was available to all women, and none of the women declined to participate.

TABLE 1—Variables Studied

Demographic	Physiological	Satisfaction	Cost
Ethnicity	Maternal	Patient satisfaction tool	Personnel
Age	Weight gain		Hours delivering prenatal care
Education	Hemoglobin		Consultation (nurse-based clinic)
Medicaid	Complications at the time of delivery		No. of appointments made/kept
Marital status	Neonatal		
Gravida	Weeks of gestation		
No. of prenatal visits	Birth weight		
Adequacy of prenatal visits	Apgar score		
Prenatal classes	Admission to Neonatal Intensive Care Unit		

Determining Pregnancy Outcomes and Satisfaction

Four broad categories of patient variables were analyzed: demographic, physiological, satisfaction with care, and cost (see Table 1). The physiological variables related to low-risk prenatal care were gathered from previous prenatal research studies.[5,6,8] The Kessner Index[9] was used to assess the adequacy of prenatal visits. This index categorizes prenatal care as adequate, intermediate, or inadequate depending on the weeks of gestation and the number of prenatal visits. Differences in the number of laboratory tests among the clinics could not be determined, as one clinic denied access to its patient records. However, both the MSC and RNC nurses used similar screening protocols and had similar guidelines for physician consultation.

Maternal satisfaction with access to care was assessed via a patient satisfaction tool[10] (PST) consisting of 27 statements that the subject rated on a 6-point Likert scale. The statements addressed five categories of satisfaction: accessibility, affordability, availability, acceptability, and accommodation. Five scores, one for each category, were generated by the PST. The readability index of the PST was at the fifth-grade level, and the internal consistency reliability coefficient was reported as .83. Both an English and a Spanish version of the tool were available to the subject. The interview took approximately 20 minutes to complete.

Determining Costs

The financial officer of each clinic provided data on five variables for study: number of staff, hourly wages (see Table 2), number of prenatal appointments made and kept, and number of hours spent delivering prenatal care. The RNC nurses kept a flowchart of the number of times they conferred with a physician and the time involved in order to determine the cost of consultation. The cost of the consultation time was computed on the basis of the hourly salary for the MSC contract physician. No physical facility costs were included in the analysis, as they were not considered germane to determining the clinics' staffing model costs.

Cost per clinic visit was determined by dividing the personnel costs by the number of kept appointments. Clinic productivity was determined by dividing the number of patients seen for prenatal care by the number of hours spent delivering prenatal care. In a cost-effectiveness analysis, the differences among the clinic costs and appointment outcomes are calculated using percentage differences. The staffing model with the lowest outcome value or cost item is identified, and then the percentage differences between this model and the other two models are calculated.[11]

Results

Demographic Variables

Subjects at the three clinics did not differ significantly on any of the demographic variables except ethnicity, years of education, and Medicaid coverage (see Table 3). Compared with the MDC and MSC, subjects at the RNC were less likely to be Hispanic, more likely to have graduated from high school, and more likely to be on Medicaid. These significant variables were further evaluated using χ^2 analysis to assess their relationship to the neonatal outcome variables. None of the χ^2 tests were significant.

Analyses of variance (ANOVAs) were computed to assess the relationship between patient ethnicity, educational level, and Medicaid coverage with satisfaction as measured by the PST. Ethnicity

TABLE 2—Clinic Staffing and Hourly Salaries

	Clinic					
	MDC		MSC		RNC	
	No. of Personnel	Hourly Salary, $	No. of Personnel	Hourly Salary, $	No. of Personnel	Hourly Salary, $
Physician						
Family practice	3	$41.96	...		...	
Contract	1	$65.00	1	35.00	...	
Obstetrical resident for consultation	...		...		1	35.00
Nurse practitioner	...		1	15.48	...	
Clinical nurse specialist	...		...		3	20.04
Registered nurse	...		5	13.53	...	
Licensed vocational nurse	3	$ 8.45	...		...	
Nurse aide	...		3	7.06	...	
Social worker	...		1	11.24	...	
Caseworker	1[a]	$ 5.19	...		...	
Clerk	1[a]	$ 5.35	1[a]	7.11	1[a]	6.47

Note. MDC = physician-based clinic; MSC = mixed-staffing clinic; RNC = nurse-based clinic.
[a]Part-time.

and Medicaid coverage were not significantly related to any of the PST statements.

Educational level was statistically significant for two PST statements ("I saw the same physician/registered nurse for all my prenatal care" and "The fees at the clinic were more than I wanted to pay"). These two statements were further analyzed using a Student Newman-Keuls post hoc *t* test to investigate pairwise differences. For the first statement ("saw the same physician/registered nurse"), those subjects with 12 years of education differed significantly from the group with 8 to 11 years of education. Cross tabulation, controlling for clinic, was done for each group; the 19 subjects who responded "never" to the statement were from the MSC. A Student Newman-Keuls test for the statement referring to the fees being too high showed that all educational levels differed significantly from the group with 0 to 7 years of education. The lower educational group was the only one in which all members (n = 4) indicated that the fees charged were "never" fair. Each clinic was represented in this group. Ninety-four percent (n = 147) of the subjects responded that the fees were "always" fair.

Analysis of the subject's reported cost for her prenatal care was done using an ANOVA. The ANOVA was significant, $F(2, 156) = 28$, $P = .000$. The grand mean for the three clinics was $29.40; the means for the individual clinics were $59.54 for MDC, $8.71 for MSC, and $19.16 for RNC. Fifty-six percent of the subjects (n = 88) reported that they paid nothing. A Pearson's *r* showed no significant relationship between subject-reported costs and the number of prenatal visits. Also, there was no significant relationship between the clinics' charge to each subject for her prenatal care, the maternal/neonatal physiological variables, and NICU admission. These differences in subject demographic data, therefore, did not appear to be related to the categorical dependent variables in the study.

Physiological Outcomes

There were no significant differences among the clinics for any of the maternal physiological variables. Twenty-eight percent (n = 20) of the subjects had an inadequate number of prenatal visits; the mean hemoglobin level at the time of hospital admission for delivery was 12 g, the mean weight gain was 30 lb, and there were no documented maternal complications at the time of delivery.

Babies of the subjects of the three clinics did not differ on any of the five neonatal variables: 85% (n = 133) were between 38 and 42 weeks of gestation, 86% (n = 134) weighed between 2500 and 4000 g, 28% (n = 43) were admitted to the NICU, 85% (n = 132) were discharged with their mother, and there were no fetal deaths or baby deaths within the first 48 hours of life.

Maternal Satisfaction

The PST internal reliability coefficient in this study was .76. ANOVA tests revealed no significant differences among the clinics for the satisfaction categories of accessibility and affordability. A Student Newman-Keuls post hoc *t* test showed the significant differences among the mean scores for the other groupings (see Table 4).

Cost Data

Table 5 shows the cost and outcome data per clinic and Table 6 shows the percentage differences among the clinics for the supplied data. Although the MSC had the lowest personnel costs, the RNC had the lowest cost per clinic visit and the highest productivity.

Discussion

The absence of significant differences among the three clinics for the maternal–neonatal physiological variables was expected and supports other reports that nurses prepared for a specific area of health care can provide quality care.[7,12–15] The number of NICU admissions approached significant differences, with the MSC having 35% (n = 20) of its subjects' babies admitted compared with 21% (n = 11) for the MDC and 23% (n = 12) for the RNC. The number of complicated pregnancies referred to the health center by each clinic was not analyzed. It is possible that the MSC referred fewer complicated pregnancies and therefore increased its number of NICU admissions. As 35% is higher than the 28% average number of newborn NICU admissions (level 1, 2, or 3) for this hospital, this

may deserve further research.[16] However, it did not affect the conclusions of this study.

The RNC subjects indicated significantly higher satisfaction than the MDC subjects in the access-to-care dimension of accommodation and overall satisfaction. This finding supports the studies, cited by OTA,[5] that reported patients often found that nurse practitioners exhibited more interest, reduced the professional mystique of health care delivery, conveyed more information, and were more skillful than physicians when providing services that depended on communication with patients.

However, the subjects at the MDC, which had no nurse practitioners, indicated significantly greater satisfaction with availability and accommodation and had a higher total satisfaction score than the subjects at the MSC, which had one nurse practitioner. One explanation for this might be that at both the MDC and RNC, subjects had only one caregiver throughout their pregnancy. All MDC and RNC subjects were able to name their care provider, while not one MSC subject could name a single individual who either had given her care or had provided health education during her prenatal visits. Subjects at the MSC saw a physician and/or nurse practitioner at each visit; however, it was not necessarily the same individual. In addition, the MSC policy of not making specific patient appointments, with the subjects sometimes waiting several hours for an appointment, was cited by many subjects as being very difficult. Lowered subject satisfaction with the MSC was probably due to its policy regarding block appointments and limited time spent with any care provider.

Prenatal care costs were based on data provided by the clinic financial officers. Costs were based on 15 minutes per visit at the MDC and 30 minutes per visit at the RNC; for the MSC costs were based on preset times for each personnel category during each 4-hour period of clinic operation. For example, during this study the clinic clerk at the MSC was assigned only 26 hours of the scheduled 104 hours of clinic operation. In addition, this clinic was frequently open for more than 4 hours, which was not accounted for in the supplied cost data. Therefore, the MSC personnel costs and cost per visit might actually be higher and their productivity lower than reported in Table 5. The lower cost per RNC appointment was primarily due to limited use of physician time, which is the most expensive resource in delivering low-risk prenatal care.

TABLE 3—Comparison of Clinics for Significant Demographic Data

	Clinic							
	MDC		MSC		RNC			
	n	%	n	%	n	%	χ^2	*P* Value
Ethnicity							26.8	.0002
Anglo		...	1	02	9	17		
Black		...	4	08	5	10		
Hispanic	52	100	47	90	38	73		
Education, y							16.5	.04
0–7	7	13	7	13	3	06		
8–11	26	50	28	54	14	27		
12	13	25	10	19	23	44		
13+	3	06	5	10	9	17		
General Equivalency Diploma	3	06	2	04	3	06		
Medicaid							11.4	.003
Yes	14	27	24	46	31	60		
No	38	73	28	54	21	40		

Note. MDC = physician-based clinic; MSC = mixed-staffing clinic; RNC = nurse-based clinic.

TABLE 4—Student Newman-Keuls Significant PST Category Means by Clinic

Significant Category	Clinic Mean Score		
	MDC	MSC	RNC
Availability	22.3	18.5[a]	23.5
Acceptability	52.3	51.6[b]	53.9
Accommodation	36.2[b]	32.3[a]	37.9
Total PST score	124.5[b]	116.6[a]	129.1

Note. PST = patient satisfaction tool; MDC = physician-based clinic; MSC = mixed-staffing clinic; RNC = nurse-based clinic.
[a]Significantly lower score compared with MDC and RNC.
[b] Significantly lower score compared with RNC.

Discussions with the RNC staff concerning the time spent waiting for physician consultation and the physician's changes to a proposed plan of care suggest the potential value of a yearly review of clinic patient care protocols by the nurse–physician team. This review would help to reduce the nonproductive waiting time for the clinical nurse specialists.

Both the RNC and the MSC had 27% missed appointments, compared with 18% at the MDC. A high incidence of missed appointments ultimately increases clinic costs because fixed personnel expenses are expended without delivering care. The lower missed-appointment incidence at the MDC may be attributed to the effectiveness of its caseworker, who called patients with missed appointments, discussed with them the importance of their clinic visits, and then tried to reschedule them.

Qualitative data can be very useful in studies of health care delivery.[17,18] This study of staffing models found significant differences in patient satisfaction and cost-effectiveness. Qualitative data from patient interviews helped ascertain that lower satisfaction with the MSC was due in part to clinic scheduling policies. In a similar vein, differences in the number of missed appointments were due in part to an effective caseworker in the MDC. These qualitative data are critically important to field studies of health care delivery that, by their nature, lack tight experimental controls.

Conclusion

Escalating health care costs, coupled with decreasing governmental economic resources, are reducing the availability of health care for underserved populations such as pregnant women.

TABLE 5—Cost and Outcome Data for Three Clinic Staffing Models

Measure	Alternative		
	MDC	MSC	RNC
Cost			
Personnel costs	$18 965	$12 381	$13 008[a]
Outcome			
Appointments made	1695	1341	1665
Appointments kept	1397	984	1216
Hours	804	962	590
Productivity	1.7	1.02	2.06
Cost-effectiveness			
Cost per appointment	$11.19	$9.23	$7.81
Cost per kept appointment	$13.58	$12.58	$10.70

Note. MDC = physician-based clinic; MSC = mixed-staffing clinic; RNC = nurse-based clinic.
[a]Includes 6 hours of consultation time.

TABLE 6—Percentage Differences in Cost and Outcome Data for Clinic Staffing Models for 3 Study Months

Measure	Clinic, %		
	MDC	MSC	RNC
Personnel costs	53	0	5
Kept appointments	42	0	24
Hours delivering Prenatal Care	35	63	0
Productivity	67	0	101
Cost per kept appointment	27	18	0

Note. MDC = physician-based clinic; MSC = mixed-staffing clinic; RNC = nurse-based clinic.

This study found that increasing the availability of low-risk prenatal care professionals through the use of nonphysician maternal health providers, with physicians available for consultation, might substantially reduce the cost of providing this care while maintaining quality. Therefore, such a system might save valuable resources. The results have direct public policy implications for the staffing of low-risk prenatal care clinics. The findings regarding the mixed-staffing model underscore the importance of analyzing environmental issues such as clinic policies regarding individual patient appointments and lack of continuity of care. □

Acknowledgments

This paper was presented at the Southern Nursing Research Conference, Atlanta, Georgia, 1990.

References

1. Lewin ME. Financing care for the poor and underinsured: an overview. In Lewin ME, ed. *The Health Policy Agenda.* Washington, DC. American Enterprise Institute for Public Policy Research; 1985:26–43.
2. Moore TR, Origel W, Key TC, Resnick R. The perinatal and economic impact of prenatal care in a low-socioeconomic population. *Am J Obstet Gynecol.* 1986;154:29–33.
3. Scholl TO, Miller LK, Salmon RW, Cofsky MC, Shearer J. Prenatal care adequacy and the outcome of adolescent pregnancy: effects on weight gain, preterm delivery, and birth weight. *Obstet Gynecol.* 1987;69:312–316.
4. Showstack JA, Budetti PP, Minkler D. Factors associated with birthweight: an exploration of the roles of prenatal care and length of gestation. *Am J Public Health.* 1984;74:1003–1008.
5. Office of Technology Assessment. *Healthy Children Investing in the Future.* Washington, DC: Office of Technology Assessment; 1988.
6. National Commission to Prevent Infant Mortality. *1985 Indirect Costs of Infant Mortality and Low Birthweight.* Washington, DC; National Commission to Prevent Infant Mortality; 1988.
7. University of Texas Health Science Center, San Antonio School of Nursing. *Developments in Prenatal Care.* San Antonio, Tex: University of Texas Health Science Center, San Antonio School of Nursing.
8. Williams RL, Binkin NJ, Clingman EJ. Pregnancy outcomes among Spanish-surname women in California. *Am J Public Health.* 1986;76:387–391.
9. Kessner DM, Kalk CE. Infant death: An analysis of maternal risk and health care. In: *Contrast in Health Status.* Washington, DC: National Academy of Sciences; Institute of Medicine; 1973;1.
10. Reed SB, Draper SJ. *The Mothers of the Nurses Prenatal Clinic (1978–1982).* Unpublished manuscript, University of Texas Health Science Center San Antonio School of Nursing.
11. Reynolds J, Gaspari KC. *Operations Research Methods: Cost-effectiveness Analysis.* Chevy Chase, Md: Primary-Health Care Operations Research; 1985. Monograph Series: Methods Paper 2.
12. Watkins LO, Wagner EH. Nurse practitioner and physician adherence to standing orders criteria for consultation or referral. *Am J Public Health.* 1982;72:22–29.
13. Hastings GE, Vick L, Lee G, Sasmor L, Natiello TA, Sanders JH. Nurse practitioners in a jailhouse clinic. *Med Care.* 1980; 18:731–744.
14. Bibb BN. Comparing nurse practitioners and physicians: a simulation study on process of care. *Eval Health Professions.* 1982;5:29–42.
15. Denton FT, Spencer BG, Gafni A, Stoddart GL. Potential savings from the adoption of nurse practitioner technology in the Canadian health care system. *Socio-Economic Plan Sci.* 1983;17:199–209.
16. Smith C. Saving newborn lives—otherwise lost. *San Antonio Express-News Sunday Magazine.* May 1, 1988;18–20, 23, 26–29, 33.
17. Lincoln Y, Guba E. *Naturalistic Inquiry.* Beverly Hills, Calif: Sage; 1985.
18. Jacob E. Clarifying qualitative research: a focus on traditions. *Educ Researcher.* 1988;17(1):16–24.

Medicaid Prenatal Care: A Comparison of Use and Outcomes in Fee-for-Service and Managed Care

James W. Krieger, MD, MPH, Frederick A. Connell, MD, MPH, and James P. LoGerfo, MD, MPH

ABSTRACT

Background. To control rising costs, state Medicaid agencies are enrolling recipients in managed care health plans (MCPs). We performed this study to assess this policy's impact on accessibility and outcomes of Medicaid-funded prenatal care.

Methods. We performed a retrospective, controlled study with three cohorts: a study group of 1106 Medicaid recipients enrolled in three MCPs, a matched comparison group of 4830 recipients receiving care in the fee-for-service (FFS) system, and a second matched comparison group of 4434 non-Medicaid enrollees of the same MCPs. Data on prenatal care use and birth outcomes were obtained through linkage of claims and discharge files with birth certificate files.

Results. Medicaid recipients enrolled in MCPs used prenatal care similarly to those in the FFS system and showed equal or modestly improved birth-weight distributions. However, Medicaid MCP enrollees showed poorer use of prenatal care and birth outcomes compared with non-Medicaid enrollees of the same plans.

Conclusions. Enrollment in MCPs has a neutral or small beneficial effect on the prenatal care received by the Medicaid population. However, providing financial access and modifying the system of care for this population did not result in parity with the general population. (*Am J Public Health.* 1992;82:185–190)

Introduction

Concern about access to health care, particularly among uninsured Americans, has been in the forefront of health policy discussions in recent years.[1–5] Diminishing access to prenatal care and a slowing of the rate of decline of infant mortality have focused attention on obstetrical care.[6,7] Numerous states have expanded their Medicaid programs in efforts to address these issues. In the early 1990s, Medicaid will cover approximately one quarter of all births.[8] This expansion of Medicaid will increase the already substantial costs of the program, rising at an annual rate of 6.2% since 1968 and now accounting for 36.4% of all state health spending.[9]

Many state Medicaid programs have initiated managed care systems in efforts to increase access while controlling costs. In these systems, the Medicaid agency contracts with a group of providers who deliver specified services to enrolled Medicaid recipients, generally in return for capitated payments.[10] Recipients may receive care only from participating providers. Medicaid managed care enrollment has increased from 187 340 in 1981 to 2 837 500 in 1991, and this growth is expected to continue.[9,11–14] Of those enrolled in managed care, 36% are health maintenance organization (HMO) members and 45% are in primary care case management fee-for-service programs. Approximately 11% of all Medicaid recipients are currently enrolled in managed care programs. Most of the enrollees are young women and children eligible through the Aid to Families with Dependent Children (AFDC) category, and obstetrical care accounts for a large proportion (33% to 41%) of Medicaid-AFDC expenditures.[15]

Available evidence indicates that Medicaid managed care may change use patterns and generate cost savings of 2% to 15% relative to fee-for-service care.[12,16–19] Less is known about the effects of managed care on accessibility and outcomes of care for Medicaid and low-income populations. Some observers have expressed concern that mandatory enrollment in designated health plans may jeopardize the quality of care received, while others suggest that managed care may improve access and continuity of care.[20–23] The Medicaid competition demonstration evaluation found no differences in the timing of initiation of prenatal care or in the distribution of birth weights between members of two mandatory enrollment managed care programs and the general AFDC population.[18] A recent study limited to Medicaid managed care members receiving care at one institution reached similar conclusions.[24]

In light of the likely continued expansion of Medicaid managed care and the large proportion of such care accounted for by maternity services, we performed this study of the impact of managed care on prenatal care for the Medicaid popula-

James W. Krieger is with the Department of Medicine, School of Medicine; Frederick A. Connell is with the Department of Health Services, School of Public Health; and James P. LoGerfo is with the Departments of Medicine and Health Services, Schools of Medicine and Public Health—all at the University of Washington.

Requests for reprints should be sent to James W. Krieger, MD, MPH, Division of General Internal Medicine, Harborview Medical Center, ZA-60, 325 Ninth Avenue, Seattle, WA 98104.

This paper was submitted to the journal November 6, 1990, and accepted with revisions September 13, 1991.

tion in Washington State. The study addresses two issues. First, does enrollment of pregnant Medicaid recipients in managed care plans result in different use and outcomes of prenatal care relative to Medicaid recipients receiving care in the fee-for-service system? Second, does managed care enrollment further the goal of promoting equity in prenatal care use and outcomes between Medicaid and non-Medicaid populations?

Methods

We conducted a retrospective, controlled study involving three cohorts: (1) women whose care was financed by Medicaid and provided by a managed care plan (the Medicaid managed care study group), (2) women whose care was financed by Medicaid and provided by the fee-for-service sector (the Medicaid fee-for-service comparison group), and (3) women whose care was financed by private enrollment payments and provided by a managed care plan (the non-Medicaid managed care comparison group).

Subject Selection

To select the two managed care groups, we used computerized discharge records to identify all women who delivered live infants between July 1983 and September 1988 while enrolled in three of the four managed care plans in Washington State that enrolled Medicaid recipients. (One of the four plans was unable to participate because of difficulties with its data system. One of the three participating plans declined to provide data on its non-Medicaid enrollees.) These records indicated the source of payment for each woman (Medicaid or non-Medicaid). To select the Medicaid fee-for-service group, we extracted from the Washington Medicaid Management Information System all hospital claims records of women who delivered live infants between February 1985 and January 1988 and who lived in counties in which the three managed care plans were operating. We excluded multiple births and stillbirths from all cohorts because of their distorting effects on birth weight.

We used data from both state Medicaid eligibility files and managed care plan enrollment files to characterize the system of care used by each individual during each month of pregnancy, that is, whether a woman (1) was enrolled in a managed care plan with either Medicaid or private financing, (2) was enrolled in fee-for-service Medicaid, or (3) had no identified system of care. Of the 1552 women identified from the managed care plans who had any months of Medicaid funding, 19% also had periods of private financing. To avoid misclassification, these women with mixed enrollment were excluded.

The length and timing of enrollment of each woman determines, in part, the degree of influence the source of care has on her pregnancy. Therefore, we determined the initiation and duration of each woman's enrollment in her system of care. Because nearly all (94%) women were continuously enrolled, we used the number of months enrolled to control for the degree of exposure to the system of care.

To control for local differences in Medicaid populations and for local variations in use of prenatal services, we frequency-matched by county of residence nine non-Medicaid managed care and nine Medicaid fee-for-service controls to each of the Medicaid managed care subjects in two plans. (For the remaining plan, no non-Medicaid managed care controls were available, and the limited number of available Medicaid fee-for-service controls permitted matching only one control for every two Medicaid managed care subjects. Overall, we were able to include approximately four controls for each case.)

Data Linkage

We used the Washington State birth certificate files as the source of dependent and confounding variables for all subjects. A computerized algorithm was created to find a unique birth certificate that matched each claim or discharge record. (Details of the linking process are available from the authors.) More than 90% of the claims and discharge records were linked with a unique birth certificate. We manually validated the accuracy of record linkages for all 1106 Medicaid managed care subjects and for a 10% random sample of the two comparison group subjects, and found 1.7% of the matches in error. The unlinked records did not differ significantly from the linked records with respect to maternal age, race, or enrollment history.

Analysis

We assessed the adequacy of the process of prenatal care with three indicators, each evaluating different aspects of care. To measure inadequate initiation of care, we computed the percentage of women whose care began in the third trimester or who received no care at all (late or no prenatal care).[25,26]

To measure the adequacy of use of care *once care was initiated,* we employed an expected visit index developed by Kotelchuck[27]: the ratio of the actual number of prenatal visits to an expected number, based on standards developed by the American College of Obstetricians and Gynecologists.[28] Use is inadequate if fewer than 50% of the expected visits occur.

Third, we computed a modification of the standard Kessner Index[29] that takes into account the current American College of Obstetricians and Gynecologists standards (available upon request). It establishes an optimal standard of care combining both initiation and use and does not adjust use for late initiation.

The outcome of prenatal care was measured by the percentage of low-birthweight (<2500 g) infants in each group.[25,30]

For univariate analyses, we tested for differences between proportions by calculating an odds ratio (OR) and its 95% confidence interval (CI).[31,32]

Because of the potential for selection bias (the Medicaid beneficiaries chose whether to receive care from fee-for-service or managed care providers), we controlled for factors other than source of care that may have affected the use of prenatal care and birth weight in two ways. First, as explained above, we frequency-matched by county. Second, we used multivariate logistic regression to control for potential confounding by age, race, smoking, parity, marital status, interbirth interval, prior preterm deliveries, and length of enrollment, as these have all been associated with the use of prenatal care and/or low birth weight.[25,33-38] We structured regression equations in a standard fashion.[39] In the final model, an OR was considered significantly different from one if its 95% CI did not include one.

Results

Medicaid Managed Care Compared with Medicaid Fee for Service

We first present the results of the comparison of Medicaid recipients enrolled in managed care plans and Medicaid recipients receiving care in the fee-for-service sector.

In general, the Medicaid beneficiaries in the managed care and fee-for-service systems of care were similar with respect to age, months of enrollment in Medicaid, race, marital status, smoking, parity, and prior preterm deliveries. They differed in that the Medicaid fee-for-ser-

vice group had a significantly shorter average interbirth interval than the managed care cohort (Table 1). Shorter interbirth intervals are associated with lower birthweight infants, but only when less than 18 months.[37] The managed care members enrolled in their system of care later than those in the fee-for-service cohort: 25.3% of the former enrolled in the third trimester compared with 15.2% of the latter ($P < .001$).

Differences in the use and outcomes of prenatal care between the managed care and fee-for-service groups varied among the three managed care plans. Therefore, we analyzed the data stratified by plan and cannot present aggregate results for all managed care plans. All ORs in Table 2 express the risk of inadequate care in the managed care group relative to the fee-for-service group. Adjustment for months of enrollment, maternal age, and maternal race did not greatly change these ORs. Also, additional adjustment for marital status, parity, smoking, number of prior preterm pregnancies, and interbirth interval did not further change the ORs.

Two of the managed care plans had fewer women with late or no prenatal care relative to fee-for-service care (only one significantly so). Members of the remaining plan received significantly later care than their fee-for-service counterparts. This finding did not change when we confined the analysis to the subset of women who were enrolled for all 9 months of pregnancy (38% of the total sample of Medicaid subjects). Any impact of the source of care would be most unambiguously demonstrated in this group.

For all three plans, we found no important or significant differences in the expected visit index between women in the managed care and fee-for-service systems. Thus, once a woman initiated care, it appeared that she received a similar number of visits in both systems (Table 2). In the subset of women enrolled for 9 months, a nonsignificant trend showing fewer women with inadequate numbers of visits in the managed care system emerged (adjusted ORs ranged from 0.41 to 1.00).

The modified Kessner Index indicated that two managed care plans had fewer women with overall inadequate care (one significantly so), while the remaining plan had significantly more (Table 2). Restricting the analysis to those women enrolled for all 9 months of pregnancy did not substantially alter these results.

Two of the managed care plans showed improved outcomes, as measured by the proportion of low-birth-weight infants (one significantly so). The remaining plan showed a nonsignificant trend toward poorer outcomes (Table 2). Restricting the analysis to those women enrolled for all 9 months of pregnancy did not importantly alter these results.

TABLE 1—Description of Study Groups by Source of Care

Descriptor	Medicaid Managed Care	Medicaid Fee for Service	Non-Medicaid Managed Care
Total subjects	1108	4435	4820
Mean age, y	23.9	23.8	28.3*
Mean months enrolled during pregnancy	6.2	6.7*	8.4*
Maternal race, %			
White	76.1	71.5	87.1*
Black	17.4	17.4	4.3*
Other	6.5	11.5	8.2
Smokers, %	54.3	56.3	13.6*
Unmarried, %	74.2	81.3**	12.7*
Mean no. of prior pregnancies	1.9	1.7***	1.3*
Mean no. of preterm pregnancies	0.47	0.40****	0.30*
Mean interval between births, mo	32.4	20.5*	28.2*

Significant difference between Medicaid managed care group and comparison group:
*$P < .001$. ***$P < .01$.
**$P < .005$. ****$P < .05$.

Medicaid Managed Care Compared with Non-Medicaid Managed Care

We next examined whether differences existed in the use and outcomes of prenatal care between Medicaid and non-Medicaid enrollees of managed care plans.

Members of the Medicaid group differed notably from the non-Medicaid managed care plan members. The Medicaid group was younger, had enrolled in care for fewer months of pregnancy, and had a higher prevalence of smokers, single mothers, non-Whites, prior pregnancies, and prior preterm pregnancies. While 25.3% of Medicaid managed care members enrolled during the last trimester of pregnancy, only 2.3% of the non-Medicaid members did so ($P < .001$).

Medicaid beneficiaries showed significantly poorer indices of use and birth outcomes than non-Medicaid enrollees. We present aggregated data from two plans because interaction between plan and source of care was not significant. As the remaining plan did not provide data on its non-Medicaid enrollees, it is not included in this analysis.

Medicaid enrollees experienced significantly higher rates of late or no prenatal care, an inadequate visit index, inadequate overall care as measured by the modified Kessner Index, and low-birthweight infants. In Table 3, the ORs indicate the risk of poor-quality care among Medicaid enrollees relative to non-Medicaid enrollees.

The unadjusted ORs reflect the experience of the Medicaid population (which is younger and has a higher proportion of Black and single mothers) relative to the non-Medicaid group. This is the appropriate measure to use when examining whether managed care can compensate for these risk factors of poor access and outcomes and actually bring the Medicaid group to the same level as the non-Medicaid group. The managed care plans were unable to do so.

The adjusted ORs address the question of whether there were differences in use and outcomes if the two groups had equivalent risk factors (age, race, marital status). After adjustment, Medicaid members continued to receive inadequate care at higher rates but no longer differed significantly in the proportion of low-birthweight infants.

Additional adjustment for plan, parity, interpregnancy interval, prior preterm births, and smoking did not cause further changes in the ORs.

Discussion

Enrollment of Medicaid beneficiaries in three managed care plans in Washington State was, on balance, associated with equal use of prenatal care relative to that of women served by the fee-for-service

TABLE 2—Medicaid Managed Care (MMC) vs Medicaid Fee for Service (FFS): Risk of Inadequate Prenatal Care and Low Birth Weight by Plan

Index	Plan 1			Plan 2			Plan 3		
	FFS	MMC	OR[a] (95% CI)	FFS	MMC	OR[a] (95% CI)	FFS	MMC	OR[a] (95% CI)
No. of subjects	2785	310		1647	183		338	618	
Late or no prenatal care, %	18.9	29.3	1.92 (1.45, 2.63)	21.1	13.9	0.59 (0.32, 1.07)	15.9	12.7	0.69 (0.48, 0.99)
Inadequate expected visit index, %	7.2	5.9	0.80 (0.45, 1.42)	4.5	5.8	1.56 (0.69, 3.57)	12.3	13.3	1.03 (0.68, 1.54)
Inadequate modified Kessner Index, %	23.1	32.3	1.75 (1.32, 2.38)	30.1	16.9	0.49 (0.27, 0.87)	20.8	19.0	0.84 (0.58, 1.12)
Low birth weight, %	8.4	4.6	0.40 (0.21, 0.78)	5.6	7.3	1.15 (0.52, 2.57)	7.1	8.0	0.83 (0.49, 1.25)

Note. OR =odds ratio; CI = confidence interval.
[a]OR (and 95% CI) for inadequate outcome in managed care relative to fee for service, adjusted for months enrolled, maternal age, and maternal race.

TABLE 3—Medicaid Managed Care (MMC) vs Non-Medicaid Managed Care (NMMC): Risk of Inadequate Prenatal Care and Low Birth Weight

Index	NMMC	MMC	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
No. of subjects[a]	4435	493		
Late or no prenatal care, %	9.9	23.6	2.83 (2.24, 3.56)	1.86 (1.39, 2.17)[b]
Inadequate expected visit index, %	1.2	5.8	5.21 (3.18, 8.54)	2.45 (1.33–4.50)[b]
Inadequate modified Kessner Index, %	10.7	26.7	3.03 (2.42, 3.81)	1.93 (1.50–2.56)[b]
Low birth weight, %	3.3	5.6	1.74 (1.14, 2.66)	0.88 (0.52–1.49)[c]

Note. OR = odds ratio; CI = confidence interval.
[a]Subjects from plans 1 and 2; plan 3 did not contribute non-Medicaid data and was excluded.
[b]OR (and 95% CI) for inadequate outcome in Medicaid vs non-Medicaid groups, adjusted for maternal race and marital status (adjusting for maternal age did not change the OR).
[c]Adjusted for maternal age, maternal race, and marital status.

sector and equal or modestly improved rates of low birth weight. These findings are consistent with those of previous studies.[18,24] In light of the data showing modest cost savings associated with managed care enrollment,[12,16–19] this policy appears promising. However, any attempt to generalize about the potential of managed care to improve birth outcomes and access to prenatal care must be tempered by the recognition that its impact differed within each of the managed care plans. Medicaid managers must carefully consider each plan in which they enroll recipients and monitor quality of care on a continuous basis.

Enrollment of the Medicaid beneficiaries in "mainstream" managed care did not eliminate the large gap in prenatal care use and birth outcomes between the Medicaid population and the general population of managed care enrollees. More Medicaid enrollees delivered low-birth-weight infants and more had inadequate access to prenatal care than general plan members. Even after adjusting for several risk factors, Medicaid members showed poorer use of care. Simply ensuring the availability of providers or extending eligibility is not sufficient to equalize the rates of adequate prenatal care use and good birth outcomes between the Medicaid and non-Medicaid populations.[18,40] The structure and content of the care offered by providers seems critical in promoting improved outcomes.

Our data demonstrated late enrollment in Medicaid by pregnant women regardless of system of care. This suggests that additional efforts are needed to facilitate earlier enrollment, such as improvement of the capacity of the maternity care system, public education, case finding, and enhanced social support services.[26]

We noted generally equal numbers of prenatal visits among Medicaid managed care and fee-for-service users once care was initiated, as reflected in the expected visit index. This finding agrees with several studies that found equal use of ambulatory services (not specifically prenatal care) among Medicaid beneficiaries in managed care and fee-for-service systems of care[17,41,42] while differing from others showing decreased[19,43,44] or increased[16] use of outpatient services in managed care systems. Several studies of non-Medicaid populations have compared the quality of maternity care received by HMO members relative to fee-for-service users. Their results are consistent with our findings of equal or better care in the managed care setting.[45–52]

There was not a consistent association of earlier initiation of prenatal care and lower rates of low birth weight. Among members of plan 1, while more women in the Medicaid managed care group had late or no prenatal care, they

also had fewer low-birth-weight infants than their fee-for-service counterparts. Medicaid recipients enrolled in plan 2 showed no significant differences in initiation of care and low-birth-weight rates relative to their fee-for-service counterparts. Recipients in plan 3 had improved timing of initiation of care (with borderline statistical significance) but equal rates of low birth weight.

These findings may be explained by noting that the association between prenatal care and birth outcomes is best studied with the individual as the unit of analysis, rather than the group. Consistent with this hypothesis is the study by Quick et al.[51] that found that members of the Kaiser-Permanente HMO in Oregon started care significantly later and received fewer visits but had modestly heavier infants relative to the general population. Despite this inverse relationship between use and outcomes, Quick et al. found an overall association of prenatal care with heavier infants in their study population when using the individual as the unit of analysis. This may reflect a "start-up" phenomenon among new users of managed care in which delays occur in establishing contact with a primary provider because of the more complex organizational structure of managed care plans. Or perhaps the women who chose managed care plans had other associated characteristics that were related to lower use of prenatal care yet better birth outcomes.

The most important caveat in interpreting the results of this study is that we were probably unable to control completely for the confounding introduced by selection bias. The potential for confounding by selection bias was strongest in the comparison of the Medicaid managed care and fee-for-service cohorts. However, eligibility for Medicaid in itself ensures a great deal of economic and social homogeneity. Additionally, we were able to control for the potential confounding effects of many covariates using available data from the birth certificates. The characteristics of the two Medicaid groups (Table 1) offer some evidence that the two cohorts are, in fact, similar.

Although several authors have questioned the reliability of some elements of birth certificate data,[25,29,53] the data we used in this study appear to be reasonably accurate. P. Starzyck (personal communication, August 1989) compared information from birth certificates with data in medical records for 765 births taking place in Washington State. The certificate and record disagreed in 3.4% of entries for birth weight, 4.2% for month of initiation, and 9.3% for number of visits. Furthermore, we had no reason to suspect a differential miscoding or error rate in one of our study groups relative to the others.

Failure to link all claims to birth certificates might have introduced ascertainment bias into this study. We fell short (by less than 10%) of linking all claims and discharge records to birth certificates. Because the linkage rates were similar among the different sources of care, introduction of bias was unlikely.

Our sample sizes for some comparisons were smaller than we had anticipated, primarily because the necessity of conducting a stratified analysis prevented us from pooling the data from the three plans. The small sample sizes in the stratified analysis raise the possibility that the lack of differences between some managed care and fee-for-service providers for some dependent variables may have been due to a type II error. For most comparisons, our sample size was large enough to detect a 50% difference between fee for service and managed care but too small to find reliably a significant difference of 25% when setting α equal to 0.05 and β equal to 0.80. The comparison of Medicaid managed care enrollees and their non-Medicaid counterparts used a larger sample (pooled across plans) and detected significant differences in most of the dependent variables.

In conclusion, managed care may offer an opportunity for modest cost savings while promoting modest improvements in prenatal care, or at least doing no harm. However, managed care alone will not substantially narrow the differences in prenatal care access and birth outcomes between the Medicaid and general populations. Facilitating earlier enrollment in Medicaid programs may increase the proportion of low-income women with timely initiation of prenatal care. Designing prenatal care programs whose content addresses the specific needs of Medicaid beneficiaries may also narrow the gap. □

Acknowledgments

This research was supported in part by the Robert Wood Johnson Foundation Clinical Scholars Program and a grant from the Bureau of Maternal and Child Health and Resources Development (MCJ 9043).

This paper was presented in part at the annual meeting of the Association for Health Services Research, June 1989.

We acknowledge the assistance of all the organizations participating in this study: the Office of Analysis and Medical Review of the Division of Medical Assistance, Department of Health and Social Services, State of Washington (Richard Boyeson, Thomas Kern, David Langenes, James Peterson); the Group Health Cooperative of Puget Sound (Eugene Hart, Peggy Schnell, Edward Wagner); the Kaiser Permanente Center for Health Research (Jonathan Brown); and the Kitsap Physicians Service (Cindy Rice, Walter Ronish, Thomas Wegeleben). Theodore Grichuhin assisted with computer programming and data linkage. Dr. Thomas Koepsell offered significant methodological advice.

References

1. Bazzoli GJ. Health care for the indigent: overview of critical issues. *Health Services Res.* 1986;21:353–393.
2. Freeman HE, Blendon RJ, Aiken LH, Sudman S, Mullinix CF, Carey CR. Americans report on the access to health care. *Health Aff.* 1987;6(1):6–18.
3. Newacheck PW. Access to ambulatory care for poor persons. *Health Services Res.* 1988;23:401–419.
4. Akin BV, Rucker L, Hubbell FA, Cygan RW, Waitzkin H. Access to medical care in a medically indigent population. *J Gen Intern Med.* 1989;4:216–220.
5. Woolhandler S, Himmelsten DU. Resolving the cost/access conflict: the case for a national health program. *J Gen Intern Med.* 1989;4:54–60.
6. Hughes D. *The Health of America's Children.* Washington, D.C.: Children's Defense Fund; 1989.
7. Kleinman J. The slowdown in the infant mortality decline. *Pediatric Perinat Epidemiol.* 1990;4:373–401.
8. Zimmerman M, Linz JR, Brecht MC, Dooley K. Medicaid *States Expand Coverage for Pregnant Women, Infants and Children.* Washington, DC: US General Accounting Office; 1989. GAO/HRD-89-90.
9. Congressional Research Service. *Medicaid Source Book: Background Data and Analysis.* Washington, DC: US Government Printing Office; 1988.
10. Hurley RE, Freund DA. A typology of Medicaid managed care. *Med Care.* 1988;26:764–774.
11. Freund DA, Neuschler E. Overview of Medicaid capitation and case management initiatives. *Health Care Financing Rev.* 1986 (annual suppl)21–30.
12. Freund DA, Hurley RE. Managed care in Medicaid. *Ann Rev Public Health.* 1987;8:137–163.
13. Office of Medicaid Managed Care, Medicaid Bureau Health Care Financing Administration, 1991.
14. Eckholm E. Late starter in Medicaid, Arizona shows the way. *New York Times.* August 7, 1991:1.
15. Howell EM, Brown GA. Prenatal, delivery and infant care under Medicaid in three states. *Health Care Financing Rev.* 1989;10:1–15.
16. Long SH, Settle RF. An evaluation of Utah's primary care case management program for Medicaid recipients. *Med Care.* 1988;26:1021–1032.
17. Bonham GS, Barber GM. Use of health care before and during Citicare. *Med Care.* 1987;25:111–119.

18. Freund DA, Rossiter LF, Fox PD, et al. Evaluation of the Medicaid competition demonstrations. *Health Care Financing Rev.* 1989;11:81–97.
19. Hurley R, Freund DA, Taylor DE. Emergency room use and primary care case management: evidence from four Medicaid demonstration programs. *Am J Public Health.* 1989;79:843–846.
20. Rosenbaum S, Hughes D, Butler E, Howard D. Incantations in the dark: Medicaid, managed care and maternity care. *Milbank Q.* 1988;66:661–693.
21. Schwartz R. Quality of care and managed care Medicaid. Presented at the 115th Annual Meeting of the American Public Health Association, New Orleans, La, October 18–22, 1987.
22. Welch WP, Miller ME. Mandatory HMO enrollment in Medicaid: the issue of freedom of choice. *Milbank Q.* 1988;66:618–637.
23. Anderson MD, Fox PD. Lessons learned from Medicaid managed care approaches. *Health Aff.* 1987;6:71–86.
24. Goldfarb NI, Hillman AL, Eisenberg JM, et al. Impact of a mandatory Medicaid case management program on prenatal care and birth outcomes. *Med Care.* 1991;29:64–71.
25. Institute of Medicine. *Preventing Low Birthweight.* Washington, DC: National Academy of Sciences; 1985.
26. Institute of Medicine; Brown SS, ed. *Prenatal Care: Reaching Mothers, Reaching Infants.* Washington, DC: National Academy of Sciences; 1988.
27. Kotelchuck M. Overview of prenatal care utilization index. Chapel Hill, NC: School of Public Health, University of North Carolina; 1987. Mimeo.
28. American College of Obstetricians and Gynecologists. *Standards for Obstetric-Gynecologic Services.* 6th ed. Washington, DC: American College of Obstetricians and Gynecologists; 1985.
29. Kessner DM, Kalk CE, Singer J. Assessing health quality: the case for tracers. *N Engl J Med.* 1973;288:189–194.
30. American Academy of Pediatrics/American College of Obstetrics and Gynecology. Standard terminology for reporting of reproductive health statistics in the United States. *Public Health Rep.* 1988;103:464–471.
31. Kelsey JL, Thompson WD, Evans AS. *Methods in Observational Epidemiology.* New York, NY: Oxford University Press; 1986.
32. Brown CC. The validity of approximation methods for interval estimation of the odds ratio. *Am J Epidemiol.* 1981;113:479–480.
33. Comstock G, Shah SK, Meyer MB, Abbey H. Low birthweight and perinatal mortality rate related to maternal smoking and socioeconomic status. *Am J Obstet Gynecol.* 1971;111:53–60.
34. Lewis R, Charles M, Patwary K. Relationships between birthweight and selected social, environmental and medical factors. *Am J Public Health.* 1973;63:973–981.
35. Petterson F, Melander S. Prediction of birthweight: results of a multiple regression analysis. *Upsala J Med Sci.* 1975;80:135–142.
36. Gortmaker SL. The effects of prenatal care on the health of the newborn. *Am J Public Health.* 1979;69:653–660.
37. Klebanoff MA. Short interpregnancy interval and the risk of low birthweight. *Am J Public Health.* 1988;78:667–670.
38. McDonald TP, Coburn AF. Predictors of prenatal care utilization. *Soc Sci Med.* 1988;27:167–172.
39. Kleinbaum DG, Kuppen LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods.* New York, NY: Van Nostrand Reinhold; 1982.
40. Piper JM, Ray WA, Griffin MR. Effect of Medicaid eligibility expansion on prenatal care and pregnancy outcome in Tennessee. *JAMA.* 1990;264:2219–2223.
41. Gaus C, Caper B, Hirschman C. Contrasts in HMO and fee-for-service performance. *Soc Security Bull.* May 1976:3–14.
42. Johnson RE, Azevedo DJ. Comparing the medical utilization and expenditures of low income health plan enrollees with Medicaid recipients and with low income enrollees having Medicaid eligibility. *Med Care.* 1979;17:953–966.
43. Fuller NA, Patera MW, Koziol K. Medicaid utilization of services in a prepaid group practice health plan. *Med Care.* 1977;15:705–737.
44. Rabin DL, Bush PJ, Fuller NA. Drug prescription rates before and after enrollment of a Medicaid population in an HMO. *Public Health Rep.* 1978;93:16–23.
45. Shapiro S, Weiner L, Densen PM. Comparison of prematurity and perinatal mortality in a general population and in the population of a prepaid group practice medical care plan. *Am J Public Health.* 1958;48:170–187.
46. Shapiro S, Jacobzinger H, Densen PM, Weiner L. Further observations of prematurity and perinatal mortality in a general population and in the population of a prepaid group practice medical care plan. *Am J Public Health.* 1960;50:1314–1317.
47. Shapiro S, Williams JJ, Yerby AS, Densen PM, Rosner H. Patterns of medical use by the indigent aged under two systems of medical care. *Am J Public Health.* 1967;57:784–790.
48. Shapiro S, Brindle J. Serving Medicaid eligibles. *Am J Public Health.* 1969;59:635–641.
49. Newport J, Roemer M. Comparative perinatal mortality under medical care foundations and other delivery models. *Inquiry.* 1975;12:10–17.
50. Cunningham FC, Williamson JW. How does the quality of health care in HMOs compare to that in other settings? *Group Health J.* 1980;1:4–25.
51. Quick J, Greenlick M, Roghmann K. Prenatal care and pregnancy outcome in an HMO and general populations: a multivariate cohort analysis. *Am J Public Health.* 1981;71:381–390.
52. Wilner S, Schoenbaum SC, Monson RR, Winickuff RN. A comparison of the quality of maternity care between a health maintenance organization and fee-for-service practices. *N Engl J Med.* 1981;304:784–787.
53. David RJ. The quality and completeness of birthweight and gestational age data in computerized birth files. *Am J Public Health.* 1980;70:964–973.

Attitudes about Infertility Interventions among Fertile and Infertile Couples

L. Jill Halman, RN, PhD, Antonia Abbey, PhD, and Frank M. Andrews, PhD

ABSTRACT

Background. There has been marked progress in the development of infertility interventions. This paper reports attitudes about 11 interventions for infertility.

Methods. Face-to-face interviews were conducted with each member of 185 infertile and 90 presumed fertile couples in southeastern Michigan.

Results. Seven of these interventions were generally viewed favorably and four were generally viewed negatively, regardless of the couple's fertility status. Infertile couples viewed all interventions, except for adoption, more favorably than did fertile couples. Multidimensional scaling was used to cluster the interventions according to similarity in endorsement. These clusters form a continuum from interventions that allow only one member of the couple to be a biological parent to the most noninvasive techniques. All clusters remain roughly equidistant from adoption, in which neither member of the couple is a biological parent.

Conclusions. Interventions that produce a child who is biologically related to only one member of the couple were viewed most negatively. Members of couples who were receiving fertility treatment made finer discriminations among infertility interventions than did individuals who had not received treatment. (*Am J Public Health.* 1992;82:191–194)

Introduction

Over the last 40 years, there has been marked progress in the development of interventions available to couples having fertility problems. This progress has been followed by criticism and debate concerning the moral and ethical issues involved in the use of these interventions. As early as 1951, infertility interventions were the subject of debate. The use of artificial insemination with donor's sperm was criticized at that time for its possible negative effects on the husband's ability to properly raise and care for the resultant child.[1] With the board certification of infertility specialists in 1973, increased research resulted in the availability of more highly technical and varied interventions.[2] These techniques and their acronyms, such as in vitro fertilization (IVF) and gamete intrafallopian transfer (GIFT), are making their way into everyday conversation; their appropriateness, acceptability, and ethical implications are being discussed and debated by many individuals, organizations, and religious groups.

Little empirical research has been conducted to explore what the public thinks about the interventions available today and how acceptable these interventions are to them. This paper reports such data. Our findings should have relevance for those health professionals involved in treating and working with people with fertility problems.

Methods

In this cross-sectional study, married couples were asked their attitudes about 11 infertility interventions. Both husbands and wives in 275 couples in southeastern Michigan were interviewed in 1988. One hundred eighty-five couples recognized that they had a fertility problem and were actively trying to solve it. One hundred sixty-one of these couples were recruited from nominations made by infertility specialists. Nine couples were volunteers from support groups (i.e., RESOLVE and the Endometriosis Association), and a few were marriage license applicants, volunteers recruited through newspaper advertisements, and other study participants. Ninety presumed fertile couples were interviewed. Forty-two were recruited by nominations made by infertility specialists who had presumed fertile gynecologic patients within their practices and by nominations made by obstetricians/gynecologists; 33 were recruited from marriage license applicants; and a few couples were volunteers recruited through newspaper advertisements and other study participants.

All study participants met the criteria of being married, White, and middle class,* and having no previous children by either member of the couple. Infertile cou-

*Middle class was defined as having at least a high school education and an annual family income in the approximate range of $20 000 to $100 000. Our respondents had an average education of 2.5 years of college and an average family income of $45 000.

L. Jill Halman and Frank M. Andrews are with the School of Public Health, Department of Population Planning and the Institute for Social Research, University of Michigan. Antonia Abbey is with the Department of Community Medicine, Wayne State University, and the Institute for Social Research, University of Michigan.

Requests for reprints should be sent to L. Jill Halman, RN, PhD, Institute for Social Research, Ann Arbor, MI 48106-1248.

This paper was submitted to the journal June 25, 1990, and accepted with revisions August 22, 1991.

TABLE 1—Favorability toward Infertility Interventions

	Infertile[a]				Fertile			
	Wives	Husbands	Paired *t* test	Couples	Wives	Husbands	Paired *t* test	Couples[b]
Male hormones	4.7	4.3	***	4.5 (342)	4.1	3.8	*	4.0 (166)***
Artificial insemination with husband's sperm	4.8	4.3	***	4.5 (368)	4.1	4.1	NS	4.1 (176)***
Progesterone suppositories[c]	4.5	4.3	NS	4.4 (306)	4.0	3.8	NS	4.0 (153)***
Drugs to stimulate ovulation	4.6	4.1	***	4.4 (364)	3.9	3.7	NS	3.8 (166)***
Tying cervix[c]	4.4	4.0	***	4.3 (276)	3.9	3.7	NS	3.8 (142)***
In vitro fertilization	4.2	4.0	**	4.1 (363)	3.8	4.0	NS	4.0 (176)
Adoption	4.1	4.0	*	4.0 (369)	4.2	4.4	NS	4.3 (179)***
Artificial insemination with husband's and donor's sperm mixed	2.6	2.4	NS	2.5 (355)	2.2	2.4	NS	2.3 (170)*
Artificial insemination with donor's sperm	2.5	2.3	*	2.4 (369)	2.0	2.3	*	2.1 (177)***
Surrogacy (wife's ova, husband's sperm)	2.2	2.0	NS	2.2 (365)	1.9	2.3	***	2.1 (179)
Surrogacy (surrogate's ovum, husband's sperm)	1.9	2.0	NS	2.0 (367)	1.7	2.0	***	1.8 (179)

Notes. Numbers in columns are mean scores; number of cases is in parentheses. The scale ran from 5 (favorable) to 1 (unfavorable). NS = not significant.
[a]Infertile couples are defined as those couples who have seen an infertility specialist.
[b]Mean scores of infertile couples were compared with mean scores of fertile couples using the Student *t* test to determine the significance of differences.
[c]15% or more of the total respondents stated they did not know enough about these interventions to state an opinion.
*$P \leq .1$.
**$P \leq .05$.
***$P \leq .01$.

ples met the additional criteria of having seen an infertility specialist, feeling they were having a difficult time having a baby, and not yet having completed the most advanced infertility treatments (e.g., IVF). Although our study originally used trying to have a child for 12 months or longer as a criterion to define fertility problems, we found that several couples had already sought treatment from infertility specialists before 12 months had elapsed. For the purposes of this data analysis, it was believed that couples who had already been involved in discussions and decision making with their physicians about infertility interventions should be included with the other infertile couples to give a more accurate picture of attitudes about infertility interventions among couples presented with this problem. Presumed fertile couples met the additional criteria of the wife's having a regular menstrual cycle and having no known gynecological problem that would affect the couple's fertility status. Eighty one percent of the referred and eligible couples participated in this study.

One-hour face-to-face interviews were conducted separately with each member of the couple.

Individual members of couples were asked, "People have different opinions about various methods for having a child. Please tell me how you feel about the possibility of *your* using each of the following methods, *if it were needed. . . .*" The 11 infertility interventions measured include artificial insemination with the husband's sperm (AIH), artificial insemination with a donor's sperm (AID), artificial insemination using both the husband's and a donor's sperm mixed together (AIH&D), hormones used to stimulate ovulation in women, hormones used to increase the sperm count in men, progesterone (a specific hormone) vaginal suppositories used to prevent miscarriage, tying the cervix to prevent miscarriage, in vitro fertilization (IVF), adoption, surrogate mother with husband-donated sperm and wife-donated ova (S/W), and surrogate mother with husband-donated sperm and the surrogate's ovum (S/S). Attitudes were measured on 5-point Likert-type scales, with options ranging from strongly opposed to personal use of a particular intervention (1) to strongly in favor of personal use of a particular method (5).

Results

In this study we examined the interventions and procedures that the fertility-problem couples had received at the time of the interview. The most frequent intervention used for women was the prescription of medications to stimulate ovulation. Adoption was currently being or had previously been considered by 20% of all fertility-problem couples and was the most commonly mentioned intervention by men. Endometriosis was a relatively common problem for the female respondents: 19% stated that they had had or were scheduled to have surgery for it and 14% stated that they had taken or were receiving medication for it. The second most common intervention reported by men was artificial insemination with the husband's sperm. Twenty percent of the men stated that they had used or were currently using this intervention.

Attitudes about Interventions

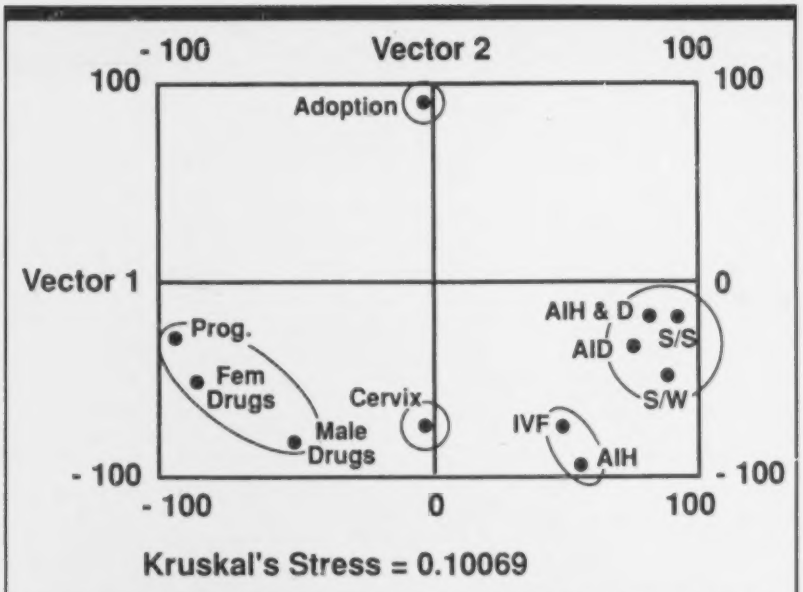
Table 1 contains respondents' attitudes about the various fertility interventions. It shows that S/S, S/W, AID, and AIH&D were viewed the most unfavorably by both fertile and infertile couples. It is noteworthy that most of these interventions allow only one member of the couple to be a known biological parent of the child. Although a couple using S/W would be known biological parents, the fetus would be carried by another woman. The remaining seven interventions are generally viewed favorably by both sets of couples. The lower acceptability of IVF among the acceptable interventions may

be due to the newness of this intervention or to the high cost, which is generally not covered by third-party payors. Infertile couples were more favorable toward all interventions, except for adoption, than were fertile couples. The Student t test shows a significant ($P \leq .01$) difference between the infertile couples' mean scores and the fertile couples' mean scores, except for IVF, AIH&D, S/W, and S/S. Adoption was the most acceptable alternative for fertile couples, whereas it was seventh for infertile couples.

As can be seen in Table 1, infertile wives and husbands showed the same hierarchical ranking of acceptability of infertility interventions, although the wives were generally more favorable toward the use of interventions. Fertile husbands and wives generally showed agreement within couples for acceptability of infertility interventions, except for male hormones, AID, S/W and S/S.

Figure 1 shows the structure of infertile couples' attitudes about 11 infertility interventions. These results were generated by applying multidimensional scaling to a matrix of correlation coefficients that reflect similarities of endorsements. In the figure, small distances between interventions imply high similarity in patterns of endorsement, and large distances imply low similarity (i.e., independence) in patterns of endorsement.** As can be seen in Figure 1, infertile couples show five clusters according to similarity of endorsement: (1) interventions in which only one member of the couple is a known biological parent, or another person is involved (S/S, S/W, AID, and AIH&D); (2) IVF and artificial insemination with husband's sperm (AIH); (3) temporary tying of the cervix to prevent miscarriage; (4) hormone medications; and (5) adoption.

Figure 2 shows multidimensional scaling results for fertile couples. Fertile couples showed less discrimination among the 11 techniques than the infertile couples did, with three clusters being depicted: (1) interventions in which only one member of the couple is a known biological parent, or another person is involved (S/S, S/W, AID, and AIH&D); (2) adoption, and (3) the remaining six interventions. All interventions, for both the infertile and fertile couples, are about equidistant from adoption. By comparing the results shown in Figures 1 and 2 with those in Table 1, we can see that 10 interventions also approximate a continuum across the diagram, moving from interventions viewed more negatively to those viewed more favorably.



AID = artificial insemination with donor's sperm; AIH = artificial insemination with husband's sperm; AIH&D = artificial insemination with husband's and donor's sperm mixed; Cervix = tying cervix to prevent miscarriage; Fem Drugs = drugs used to stimulate ovulation; IVF = in vitro fertilization; Male Drugs = drugs used to increase the sperm count; Prog = progesterone vaginal suppositories to prevent miscarriage; S/S = surrogacy using surrogate's ovum and husband's sperm; S/W = surrogacy using wife's ova and husband's sperm.

FIGURE 1—Interventions for infertile couples, multidimensional scaling.

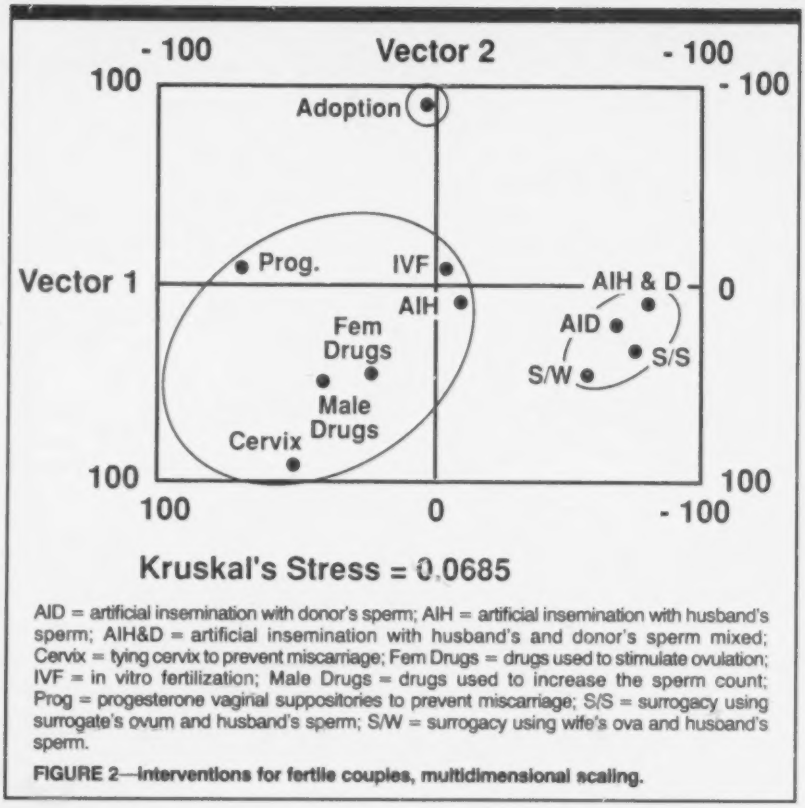
Discussion

Comparing the frequency of the procedures and interventions used among the infertile couples with the acceptability of various infertility interventions, we can see that, in the aggregate, couples are using interventions that are generally the most acceptable to them. Adoption is the one intervention that is inconsistent with this. Although adoption was frequently pursued by infertile couples, it was the least acceptable of the interventions that infertile couples felt they could use. This result may have been influenced by the respondents who participated in this study. Couples who strongly prefer to be the biological parents of their child may be the couples who seek treatment from an infertility specialist (our primary source of referral for infertile couples). Although the couples in this study may have proceeded with investigation of adoption to assure themselves of success in eventually having a child, they may not yet have been ready to accept this as their ultimate intervention.

Comparing infertile couples to fertile couples, we see that the infertile couples viewed all interventions, except for adoption, more favorably than did the fertile couples. The explanation for this may be the same as stated in the previous paragraph, but a different explanation is also possible. Some couples in our pretest who adopted after receiving infertility treatment spoke about not seriously considering adoption as an intervention initially because they had not considered it an alternative treatment. It is known that surprise and shock are among the first emotions felt by couples when they suspect they may be infertile.[3,4] It may be that knowledge of infertility comes as such a surprise to some people that they do not recognize all the options available to them and instead focus only on seeing an infertility specialist.

Both fertile and infertile couples viewed interventions in which only one member of the couple is a biological parent the most negatively. This attitude may be explained by members of the couple

**To select the appropriate dimensional solution, one should assess how well the configuration approximates the actual data. This can be measured by Kruskal's stress coefficient, which is .10 and .07 (i.e., "fair" to "good") for the two-dimensional solutions of Figures 1 and 2, respectively.



AID = artificial insemination with donor's sperm; AIH = artificial insemination with husband's sperm; AIH&D = artificial insemination with husband's and donor's sperm mixed; Cervix = tying cervix to prevent miscarriage; Fem Drugs = drugs used to stimulate ovulation; IVF = in vitro fertilization; Male Drugs = drugs used to increase the sperm count; Prog = progesterone vaginal suppositories to prevent miscarriage; S/S = surrogacy using surrogate's ovum and husband's sperm; S/W = surrogacy using wife's ova and husoand's sperm.

FIGURE 2—Interventions for fertile couples, multidimensional scaling.

preferring equity in relation to the child; that is, if both parents are unable to have a child biologically, then perhaps it is better to have neither be a biological parent and to adopt a child instead.

The multidimensional scaling shows that within their clusters, infertility interventions seem to follow a continuum from those that allow only one member of the couple to be a known biological parent to those that allow both spouses to be biological parents. These 10 interventions are roughly equidistant from adoption, presumably because neither member of the couple would be a biological parent with that intervention.

Future research is necessary to determine the extent to which the pattern of attitudes found here is representative of that in the general population. These findings have implications for health care providers regarding the reluctance that couples experiencing fertility problems may have, at least initially, to accept some interventions required for the couple to conceive. These findings could serve as a general guide for ease of acceptance of certain interventions and for the degree of emotional support some couples may require in order to make an informed decision about the use of a particular technique. Couples may benefit by having the alternative of adoption addressed and given legitimacy as an acceptable intervention from the onset of infertility treatment. The finding that infertile couples had more favorable attitudes than did fertile couples toward various interventions suggests that when couples are confronted with more information and personal necessity, their attitudes about infertility interventions become more positive. □

Acknowledgments

This work was sponsored by NICHD (grant HD21240).

We would like to express our appreciation to the physicians and nurses who took the time from their busy practices to facilitate the referral of respondents for this study. We are especially appreciative of the couples who so graciously and willingly participated in this project; without their cooperation this study would not have been possible.

A portion of these results was presented at the 1989 annual meeting of the American Psychological Association, New Orleans, La.

References

1. Lamsen HD, Pinard WJ, Meaker SR. Sociologic and psychological aspects of artificial insemination with donor sperm. *JAMA.* 1951;13:1062–1064.
2. Arel SO, Cates W. The increasing concern with infertility: why now? *JAMA.* 1983;17:2327–2331.
3. Batterman R. A comprehensive approach to treating infertility. *Health Soc Work.* 1985:46–53.
4. Menning BE. The emotional needs of infertile couples. *Fertil Steril.* 1980;34:313–319.

Nonclinical Factors Associated with Surgery Received for Treatment of Early-Stage Breast Cancer

ABSTRACT

Background. Women diagnosed with early breast cancer have had the opportunity to receive breast-conserving surgical treatment, which reduces the physical and psychological morbidity heretofore associated with breast removal.

Methods. Nonclinical factors associated with women receiving partial mastectomies with radiation (P+R) compared with modified radical mastectomies without radiation (MOR) were examined in 2238 Black and White women diagnosed, in 1985 through 1987, with early-stage breast cancer in the metropolitan Detroit area.

Results. Age at diagnosis and size of hospital were the strongest predictors of type of surgery received, with younger women (less than 55 years of age) and women treated in larger hospitals (more than 500 beds) more than twice as likely to receive P+R. Stratifying on race, age at diagnosis remained the strongest predictor for White women, followed by hospital size. Among Black women, hospital size was more strongly associated with surgery received than was age.

Conclusions. Younger women and women undergoing treatment at large hospitals are more likely to receive the breast-conserving P+R. Black women treated in small hospitals appear to be particularly unlikely to receive P+R. (*Am J Public Health.* 1992;82:195–198)

Enid R. Satariano, MPH, G. Marie Swanson, PhD, MPH, and Patricia P. Moll, PhD

Introduction

Breast cancer is the most common malignancy in women. It is estimated that 150 000 new cases are diagnosed each year in the United States.[1] Surgery remains the standard form of treatment, with over 90% of all cases receiving some surgical intervention.

The breast-sparing partial mastectomy has gained credibility in the treatment of early-stage breast cancer with the publication of the results of several randomized clinical trials undertaken in the 1970s and 1980s.[2–4] These studies demonstrated no significant difference in the duration of survival between women receiving either partial mastectomies with radiation or modified radical mastectomies without radiation for early-stage disease. With nearly 50% of all breast cancers being diagnosed at a localized stage,[5] there are clearly a substantial number of women for whom the less invasive procedure may be an option. The potential impact on the psychological and physical morbidity heretofore associated with breast removal[6,7] also is important. Laws have been enacted in a number of states requiring physicians to inform breast cancer patients of treatment alternatives. Although there has been an increase in the proportion of cases receiving partial mastectomies since 1980, the procedure has not become the standard surgical treatment for women with early-stage disease. In this report, we use data from a population-based cancer registry encompassing the racially and socioeconomically diverse, urban population of the Detroit metropolitan area to identify nonclinical factors associated with women receiving partial mastectomies with radiation compared with modified radical mastectomies without radiation in a sample of breast cancer patients with similar clinical profiles.

Methods

Cases were selected for this study who met the criteria for breast-sparing treatment.[4] The study includes 2238 Black and White women diagnosed, in 1985 through 1987, with microscopically confirmed, node-negative, incident invasive cancers of the breast less than or equal to 3.9 cm who received either partial mastectomies with radiation (P+R) or modified radical mastectomies without radiation (MOR) as the first course of treatment. More than 90% of all first primary breast cancers diagnosed during this time period in the metropolitan Detroit area were treated with one of these modalities. Cases were identified through the Metropolitan Detroit Cancer Surveillance System (MDCSS), a participant in the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program.[5] The MDCSS encompasses a population of 4.1 million residents of the metropolitan Detroit tricounty area. Medical and demographic information is obtained for incident cancer cases from all

Enid R. Satariano is with the California State Department of Health Services, Emeryville. G. Marie Swanson is with the College of Human Medicine, Michigan State University, East Lansing. Patricia P. Moll is with the School of Public Health, University of Michigan, Ann Arbor.

Requests for reprints should be sent to G. Marie Swanson, PhD, Cancer Center, Michigan State University, A-211 East Fee Hall, East Lansing, MI 48824.

This paper was submitted to the journal January 8, 1991, and accepted with revisions August 23, 1991.

TABLE 1—Surgery Received for Early Stage Breast Cancer by Selected Demographic and Hospital Characteristics: Detroit Metropolitan Area, 1985–1987

	Treatment Type			
	Partial Mastectomy with Radiation, %[a]	Modified Radical Mastectomy without Radiation, %[a]	n	Unadjusted Odds Ratios (95% Confidence Interval)
All patients	29.4	70.6	2238	
Age at diagnosis				
<55	42.4	57.6	689	2.38 (1.97, 2.88)
≥55	23.6	76.4	1549	
Race				
White	29.1	70.9	1934	0.92 (0.71, 1.19)
Black	30.9	69.1	304	
Marital status				
Married	33.0	67.0	1303	
Single	28.1	71.9	153	1.26 (0.86, 1.86)
Separated/divorced	38.3	61.7	180	0.79 (0.57, 1.11)
Widowed	19.9	80.1	533	1.98 (1.55, 2.55)
Hospital size				
≤500 beds	21.6	78.4	1248	
>500 beds	39.2	60.8	990	2.76 (2.22, 3.33)
No. of breast surgeries performed by surgeon in 1985–1987				
≤50	25.3	74.7	1475	
>50	34.0	66.0	709	1.52 (1.23, 1.84)
Year of diagnosis				
1985	31.6	68.4	694	
1986	30.9	69.1	698	1.03 (0.81, 1.30)
1987	26.2	73.8	846	1.30 (1.03, 1.63)

[a]Row percentages sum to 100%.

TABLE 2—Odds of Receiving a Partial Mastectomy with Radiation for Treatment of Early-Stage Breast Cancer by Selected Nonclinical Factors: Detroit Metropolitan Area, 1985–1987

	n	Odds Ratio[a]	95% Confidence Interval
Age at diagnosis		2.38	1.96, 3.03
<55	689		
≥55[b]	1549		
Hospital size		2.13	1.73, 2.62
≤500 beds[b]	1248		
>500 beds	990		
No. of breast surgeries performed by surgeon		1.28	1.03, 1.58
≤50[b]	1475		
>50	709		

[a]Adjusted for all other variables (all variables are adjusted for race, marital status, and year of diagnosis; age at diagnosis is also adjusted for hospital size and number of surgeries, hospital size is also adjusted for age at diagnosis and number of surgeries, and number of surgeries is also adjusted for age at diagnosis and hospital size).
[b]Referent category.

tricounty facilities that provide care for cancer patients as well as three out-of-area hospitals that provide treatment for approximately 10% of cancer patients from the surveillance area. The MDCSS is considered to have complete registration of area cancer cases, with 1.2% of the cases identified by death certificate only.

Nonclinical factors included age at diagnosis, race, marital status, year of diagnosis, size of the hospital where a patient's surgery was performed, and total number of breast cancer surgeries performed by a patient's surgeon during 1985 through 1987. Odds ratios (ORs) were calculated as a measure of association between the nonclinical factors and treatment. Stepwise logistic regression with BMDP[8] was employed to fit the data using all variables. The significance of each term was assessed by the F test. The population-attributable fraction was calculated to determine the proportion of breast-conserving treatment accounted for by the variables that were associated with this treatment.[9]

Results

Of the 2238 women included in this study, 657 (29.4%) received P+R and 1581 (70.6%) received MOR for treatment of early-stage breast cancer (Table 1). Although women 55 years of age or older account for almost 70% of the cases, proportionately fewer of them received P+R (23.6%) compared with women less than 55 (42.4%). Black women account for 13.6% of the study population. They did not differ from White women in treatment received. More women underwent surgical treatment (55.8%) in smaller hospitals (maximum of 500 beds), but they were less likely to receive P+R in these hospitals (Table 1). A majority of the women in this study (65.9%) were treated by surgeons who had performed 50 or fewer breast cancer surgeries during 1985 through 1987, yet fewer of them received P+R than those treated by surgeons who had performed more than 50 surgeries (Table 1).

The logistic regression analysis (Table 2) indicates that, controlling for the other variables, type of surgery received for early-stage breast cancer is associated with the following factors: age at diagnosis, size of hospital, and number of surgeries performed by a patient's surgeon. Younger women are more than twice as likely as older women to receive P+R (OR = 2.38; 95% confidence interval [CI] = 1.96, 3.03). Women undergoing surgery at larger institutions are twice as likely to receive P+R as women at smaller hospitals (OR = 2.13; 95% CI = 1.73, 2.62). Women treated by surgeons performing more than 50 breast cancer surgeries in the study period were slightly more likely to receive P+R than women treated by surgeons performing 50 or fewer (OR = 1.28; 95% CI = 1.03, 1.58). No significant interaction is apparent for hospital size and age at diagnosis, number of surgeries and age at diagnosis, or number of surgeries and hospital size (data not shown). On the basis of the population-attributable fraction, about 15% of the breast-conserving treatment is explained

by age at diagnosis, 16% by hospital size, and 3% by number of surgeries performed.

Since the association between specific nonclinical characteristics and treatment might not be similar for Black and White women, we stratified on race (Table 3). Age at diagnosis and hospital size remain significant for both Black and White women. Also, year of diagnosis is significant.

Age at diagnosis remains the strongest predictor for White women, while for Black women hospital size is the strongest predictor of breast-sparing treatment (Table 3). Year of diagnosis is more predictive of treatment type than the number of surgeries performed by a woman's physician for both Black and White women when the two racial groups are considered separately. The associations, however, are in the opposite direction (Table 3). This may account for the lack of year effect in the model when Blacks and Whites were pooled. A χ^2 test for trend was significant within each racial group but was stronger for White women (10.93; $P < .005$) than for Black women (4.11; $P < .05$).

TABLE 3—Odds of Receiving a Partial Mastectomy with Radiation by Selected Nonclinical Factors: Detroit Metropolitan Area, 1985–1987

	Black Women			White Women		
	n	Odds Ratio[a]	95% Confidence Interval	n	Odds Ratio[a]	95% Confidence Interval
Age at diagnosis		2.08	1.14, 3.84		2.63	2.08, 3.22
<55	121			568		
≥55[b]	183			1366		
Hospital size		3.72	1.99, 6.94		1.96	1.57, 2.45
≤500 beds[b]	159			1089		
>500 beds	145			845		
Year of diagnosis						
1985[b]	101			593		
1986	100	2.35	1.10, 5.00	598	.93	0.71, 1.21
1987	103	3.14	1.47, 6.71	743	.72	0.55, 0.94
No. of breast surgeries performed by surgeon		1.64	0.90, 2.91		1.24	0.98, 1.55
<50[b]	198			1277		
≥50	88			621		

[a]Adjusted for all other variables (all variables are adjusted for marital status; age at diagnosis is also adjusted for hospital size, year of diagnosis, and number of surgeries; hospital size is also adjusted for age at diagnosis, year of diagnosis, and number of surgeries; year of diagnosis is also adjusted for age at diagnosis, hospital size, and number of surgeries; and number of surgeries is also adjusted for age at diagnosis, hospital size, and year of diagnosis).
[b]Referent category.

Discussion

Overall, a younger woman is considerably more likely to receive P+R for treatment of early-stage breast cancer than is her older counterpart presenting with the same breast cancer. Although it has been shown in other studies[10–12] that older women typically receive less invasive (i.e., less definitive) therapy than younger women, in this study older women were less likely to receive the partial mastectomy. There are a number of possible explanations for this seeming contradiction in the age-related treatment patterns for breast cancer. Standard therapy requires that partial mastectomy be accompanied by radiation treatment. Six weeks of radiation therapy may be contraindicated by comorbid conditions or the physical status of an older woman. Physicians may be less likely to recommend this treatment for their older patients regardless of their physical status. Although many of the patients in this study would have received information on treatment options from their physicians as mandated by Michigan's breast cancer informed-consent legislation,[13] it is not known how objectively this information is presented. Older women may be more sensitive to their physician's opinions and recommendations, and thus more likely to defer to their doctor. In contrast, younger women may be more likely to read materials presented to them, to question the physician, and to take a more active role in the treatment decision-making process. Physicians may be more likely to recommend a breast-sparing technique to a younger woman, bringing in their own attitudes about "femininity" and the age relatedness of the importance of breast conservation.

The finding that women are more likely to receive P+R at a large hospital suggests that new treatment modalities may be implemented more quickly in environments associated with medical centers, research programs, and physician training, which are usually located in larger institutions. Black and White women in this metropolitan area who receive treatment at large institutions are going to the same large institutions and do not differ in their overall probability of receiving P+R compared with MOR. Black women being treated for early-stage breast cancer are less likely to receive P+R at smaller hospitals than White women being treated in hospitals of the same size. White women receiving treatment at smaller institutions are more likely to be seen in suburban hospitals, while Black women are more often seen in urban hospitals. Thus, the treatment differences in the small hospitals are urban–suburban variations. Further studies should investigate whether Black women seen in small hospitals do not receive P+R as often because their physicians are concerned about treatment compliance for the radiation component, because radiation therapy is not available at those hospitals, or because these women are not provided with adequate information about alternative treatments.

The years included in this study are those immediately following the publication of data that demonstrated no significant survival differences between women receiving P+R and MOR for treatment of early-stage breast cancer. These data have been reinforced recently with the publication of data showing similar results after 8 years of follow-up.[14] Women have the opportunity to receive a breast-conserving surgical treatment that eliminates the physical and psychological morbidity heretofore associated with breast removal. The results of this study indicate that younger women and women undergoing treatment at large hospitals are more likely to receive breast-preserving treatment. Black women being treated in small hospitals appear to be particularly unlikely to receive P+R. The time trend among White women shows a decline in the proportion receiving P+R in both 1986 and 1987, compared with 1985. This trend should be observed for the next 2 to 3 years and studied in depth if it continues. Furthermore, since no single variable considered in this study explains more than 16% of the shift toward breast-con-

serving therapy, future studies must include other factors that are likely to influence treatment decisions. Consideration should be given to comorbid conditions, factors that influence a woman's decision to undergo one form of treatment compared with another for early-stage breast cancer, information provided to women about treatment options, and physician characteristics associated with use of breast-conserving treatment. □

Acknowledgments

This work was supported in part by National Cancer Institute Contract NO1-CN-05225.

The authors wish to thank Ms. Colleen Kniffen for her careful preparation of the manuscript.

References

1. Silverberg E, Boring CC, Squires TS. *Statistics 1990.* Atlanta, Ga: American Cancer Society; 1990.
2. Veronesi U, Saccozzi R, DelVecchio M, et al. Comparing radical mastectomy with quadrantectomy, axillary dissection, and radio-therapy in patients with small cancers of the breast. *N Engl J Med.* 1981;305:6–11.
3. Hayward JL. The guy's hospital trial on breast conservation. In: Harris JR, Hellman S, Selen W, eds. *Conservative Management of Breast Cancer.* Philadelphia, Pa: JB Lippincott; 1983:77–90.
4. Fisher B, Bauer M, Margolese R, et al. Five-year results of a randomized clinical trial comparing total mastectomy and segmental mastectomy with or without radiation in the treatment of breast cancer. *N Engl J Med.* 1985;312:665–673.
5. National Cancer Institute, Division of Cancer Prevention and Control. *Annual Cancer Statistics Review.* Bethesda, Md: National Institutes of Health; 1988.
6. Bransfield DD. Breast cancer and sexual functioning: a review of the literature and implications for future research. *Int J Psychiatry Med.* 1982;12(3):197–211.
7. Steinberg MD, Juliana MA, Wise L. Psychological outcome of lumpectomy versus mastectomy in the treatment of breast cancer. *Am J Psychiatry.* 1985;142:34–39.
8. Dixon WJ, ed. *BMDP Statistical Software.* Berkeley, Calif: University of California Press; 1983.
9. Kelsey JL, Thompson WD, Evans AS. *Methods in Observational Epidemiology.* New York, NY: Oxford University Press; 1986.
10. Samet J, Hung WC, Key C, et al. Choice of cancer therapy varies with age of patient. *JAMA.* 1986;255:3385–3390.
11. Chu J, Diehr P, Feigl P, et al. The effect of age on the care of women with breast cancer in community hospitals. *J Gerontol.* 1987;42:185–190.
12. Greenfield S, Blanco DM, Elashoff RM, et al. Patterns of care related to age of breast cancer patients. *JAMA.* 1987;257:2766–2770.
13. Center for Health Promotion, Cancer and Tobacco Unit. *Breast Cancer: What You Should Know before Treatment.* Lansing, Mich: Michigan Department of Public Health; 1990.
14. Fisher B, Redmond C, Poisson R, et al. Eight-year results of a randomized clinical trial comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. *N Engl J Med.* 1989;320:822–828.

Epidemiology Summer Program to Be Held at Tufts University

The New England Epidemiology Institute and Tufts University will sponsor a 3-week summer program in epidemiology at Tufts' Medford campus July 12 to 31, 1992. The program includes both methodologic and substantive courses intended for those seeking an introduction to modern epidemiologic concepts as well as those desiring a review of recent developments in epidemiologic thinking.

Fourteen courses will be offered, including Theory & Practice of Epidemiology (introductory and advanced levels), Causal Inference, Biostatistics for Epidemiologists, Regression and Categorical Data Methods, Logistic Regression and Survival Analysis, Conducting Epidemiologic Research, The Biology and Epidemiology of Cancer, Environmental Epidemiology, Nutritional Epidemiology, Clinical Epidemiology, Concepts of Infectious Disease Epidemiology, Perinatal Epidemiology, and Epidemiology in Public Health Practice. Invited faculty include excellent teachers and prominent researchers from leading universities. Registrants may receive graduate degree credit or continuing education credit, Continuing Medical Education credit (AMA Category 1), or Certification Maintenance credits in industrial hygiene.

For more information please contact The New England Epidemiology Institute, Dept. 15, One Newton Executive Park, Newton Lower Falls, MA 02162. (Phone: 617/244-1200; FAX: 617/244-9669.)

Breast Cancer Screening Practices among Users of County-Funded Health Centers vs Women in the Entire Community

ABSTRACT

Background. Breast cancer screening rates tend to be lower among women with lower income and/or education.

Methods. Telephone surveys of random samples of women aged 50 to 75 who had visited five health centers (n = 795) and women in the same age group residing in the entire community (n = 404) were conducted in 1988.

Results. Despite the significantly lower socioeconomic level, a higher proportion of minority women, and a poorer knowledge of mammography, screening rates in the health center group were not lower than in the community sample. About half or more of the health center respondents had ever had a mammogram regardless of income, education, age, and ethnic group. Of the community respondents, 49% had ever had a mammogram, but the proportion varied significantly by income and education. Within the subgroup of women having annual incomes below $15 000, mammography use was actually higher among health center women; that is, 50% of health center vs 35% of community women reported ever having a mammogram and 31% vs 14%, respectively, reported having a mammogram in the past year.

Conclusion. The findings demonstrate the importance of publicly funded health centers in achieving screening rates among the socioeconomically disadvantaged that are comparable to (or even higher than) those in the general population. (*Am J Public Health.* 1992;82:199–203)

Dorothy S. Lane, MD, MPH, Anthony P. Polednak, PhD, and Mary Ann Burg, PhD, MSW

Introduction

In 1989 the National Cancer Institute (NCI) released a consensus statement about guidelines for mammography screening that was endorsed by 11 national organizations, including the American Cancer Society (ACS) and the American Medical Association. This statement includes a recommendation for annual mammography and breast examination for women aged 50 and older.[1] The *Guide to Clinical Preventive Services,* by the US Prevention Services Task Force, also recommends that women aged 50 to 75 have a breast examination annually and mammography annually or every 2 years.[2] And in *Healthy People: 2000* (1990), the US Department of Health and Human Services set a national goal to increase to at least 60% the proportion of women aged 50 and older who have received a clinical breast examination and a mammogram in the preceding 1 to 2 years (and to increase to at least 80% the proportion who have ever received them).[3]

Preventive health services are traditionally underused, particularly by the poor, who have competing and more urgent priorities as well as greater problems of access to all health services. And although mammography is a highly effective screening technology for breast cancer, it is not adequately used, particularly by low-income women.[3,4] Yet, despite the higher incidence of breast cancer among women of higher socioeconomic status,[5] the need for early detection of breast cancer is most evident among the socioeconomically disadvantaged because they suffer higher mortality due to later-stage diagnosis.

This paper assesses differences in breast cancer screening practices between samples of predominantly low-income women aged 50 to 75 using county-funded health centers and a population-based sample of women in the same age group residing in the towns where the health centers are located.

Methods

Background

This study was part of an NCI-supported project to increase the use of mammography and physical breast examination on a communitywide basis among women 50 years of age and older through public and physician education, as described in detail elsewhere.[6] Baseline surveys were done for needs assessment related to planning interventions to improve the use of mammography and physical breast exam, as well as to contribute to the literature currently available about factors associated with use of breast cancer screening among low-income women.

The Suffolk (NY) County Department of Health Services provides funding for five health centers that are located in the project's study towns (three centers are in the town of Brookhaven, two are in the town of Islip). Because there is no county hospital, these health centers play an important role in providing ambulatory

At the time of the study the authors were with the Department of Preventive Medicine in the School of Medicine, State University of New York at Stony Brook.

Requests for reprints should be sent to Dorothy S. Lane, MD, Department of Preventive Medicine, School of Medicine, State University of New York at Stony Brook, Stony Brook, NY 11794.

This paper was submitted to the journal May 28, 1991, and accepted with revisions October 1, 1991.

care services for low-income residents. At the time of the baseline survey, mammography services were not available at the health center sites, nor were they paid for by the county; thus, health center women were referred to private radiologists or hospitals for mammography.

Study Samples

A telephone survey was conducted in 1988 on a random sample of women aged 50 to 75, selected from computer files of women who had visited one or more of the five health centers for any reason within the prior year. Random samples were drawn to obtain 400 completed interviews in each of the two intervention towns in which the health centers were located. The sampling was proportionate to the total number of women visiting each health center in that town. Although possible telephone numbers for many women in the sample were obtained from health center records, the interviewers required additional searches. These entailed the use of standard telephone and cross-reference directories, and calls to information operators and to persons with the same surname as sampled patients. The response rate resulting from these procedures was 60.4%. Lack of a known telephone number was the most common reason for nonresponse, and the participation rate among persons with apparently correct telephone numbers was 74%. Because the proportion of women without a telephone number who had moved out of the area was unknown and we were interested only in women who were still residing in the study towns, 60.4% is an underestimate of the "true" response rate, which lies somewhere between 60% and 74%. The proportion of respondents from each of the five health centers ranged from 13% to 34%.

The same survey was also conducted in 1988 by random-digit dialing to obtain at least 500 interviews of women aged 50 to 75 residing in the entire project intervention area, which included three towns. The estimated response rate was 73.1%. Nonresidential numbers and numbers not in service were excluded. Eligible households were those with one or more women aged 50 to 75 living in the study area. Reasons why eligibility could not be determined include household's refusal to participate prior to screening for eligibility, use of continuous answering machine, and inability to speak one of the four interviewing languages. The proportion of eligibles among those screened was 29.6%. Among households that did not answer after nine attempts, a maximum of 10% of the phone numbers called were estimated to be current residential numbers based on prior experiments conducted by the contractor (Northeast Research in Orono, Me). The proportion of eligibles among those not screened for eligibility was unknown but was estimated to be the same as that among those screened. A methodological article about our community population survey provides a detailed discussion of the response rate calculation.[7] Data in this paper cover only the two towns in which the health centers are located.

Because only 5% of respondents to the community survey reported a health center as their regular source of care and 87% cited a private physician, there is little overlap between this group and the health center respondents. The two samples, however, may differ in the frequency of recent contact with a physician. The health center sample was restricted to women known to have visited the health center for any reason during the past year, including visits involving physician contacts. Unfortunately, in an effort to limit the length of the survey so it could be completed in 10 minutes, one question that was omitted from the 1988 telephone surveys was whether respondents had visited a physician in the past year. However, in a 1989 survey of women known to have visited a health center during the past year, 96% reported having seen a physician for any reason that year, compared with 83% in a random sample of community women surveyed by mail. This suggests that there probably was not a large difference between the samples in terms of recent contact with a physician.

In the survey, women were asked about their current breast cancer screening practices, their concerns that would influence their participation in screening or having a mammogram, their knowledge and attitudes related to breast cancer risk and the benefits of screening, their sources of health care and health insurance, their demographic characteristics, and their personal history of breast cancer. Those with a prior history of breast cancer were excluded from the analysis. Women were asked to rank their level of concern on a scale of 1 to 5, where 1 is "no concern" and 5 is "a great deal of concern."

Analysis

Chi-square tests were performed using an SAS mainframe program to compare various characteristics of the health center and community survey respondents. Tests with one degree of freedom (*df*) were continuity corrected. The *P* values in the tables are reported to three decimal places (as low as $P < .001$). Because of large sample sizes in some comparisons, the attainment of statistical significance may not indicate meaningful differences, and the magnitude of the differences in proportions should be considered.

Results

Sociodemographic Characteristics

As indicated in Table 1, there were significant differences between the health center and community respondents in all demographic characteristics. The health center respondents were older (43.7% aged 65 to 75 vs 31.2%) and less educated (43.0% had less than a 12th-grade education vs 20.1%); they also had a lower annual household income (52.8% made less than $15 000 vs 19.3%) and had a larger minority population (26.7% Black and Hispanic vs 6.7%). A significantly higher proportion of the health center respondents reported having a regular source of care (96.5%) compared with the community respondents (88.5%), which was statistically significant for these sample sizes ($\chi^2 = 28.744$, $df = 2$, $P < .001$).

Knowledge and Attitudes

Health center women knew less frequently than community respondents that the lifetime risk of breast cancer was about 1 in 10 (27.0% vs 33.2%; $\chi^2 = 4.577$, $df = 1$, $P < .05$), but their perception of a personal risk of breast cancer was similar to that of community respondents (likely risk = 25.0% vs 25.3%; $\chi^2 = 0.000$, $df = 1$, $P = .991$).

Health center respondents agreed strongly with statements about the benefits of mammography and early detection less frequently than did community respondents, and there was a statistically significant difference in those agreeing that there are other ways to treat breast cancer besides removing the breast and that women live longer if breast cancer is found early (Table 2).

The rank order of concerns that would influence their compliance with mammography guidelines were the same for the health center and community respondents (Table 3). A small difference between the two groups in the frequency of reporting the most common concern—the results of the mammography exam (health center = 83% vs community =

89%)—reached statistical significance (Table 3). Health center women had greater concern about cost (64%) significantly more often than did community respondents (49%) (Table 3); the level of concern with cost was also significantly higher among health center respondents (36.9% extremely concerned vs 17.1%; $\chi^2 = 74.316$, $df = 4$, $P < .001$). There was no significant difference, however, in the ranking or level of concerns about factors that might influence having a breast exam, with fear of results cited most often, followed by embarrassment.

More than 90% of women in both groups had ever heard of a mammogram, and the small difference between community women and health center women in this regard reached statistical significance (Table 4).

Screening Behavior

Despite their differences in socioeconomic status and in knowledge and attitudes, the health center and community samples differed little when asked if they had ever had a mammogram or breast exam or had had a mammogram in the past year (Table 4). Although there was some variation in mammography use in the past year by individual health centers (ranging from 23% to 41%), there was no health center where mammography use was substantially lower than it was in the community as a whole (29%). Significantly more health center women than community women reported having had a breast exam in the past year (Table 4).

Mammography use in the past year was significantly associated with income and education (and there was a trend toward declining use with increasing age) in the community sample, but there was little variation in screening use by income, educational level, or age within the health center sample (Table 5). Mammography use in the past year also did not vary significantly by race/ethnic group within the health center sample (white/non-Hispanic 28.7% [163/568], Black/non-Hispanic 28.6% [26/91], Hispanic 33.1% [40/121]; $\chi^2 = 0.945$, $df = 2$, $P = .623$); the numbers of minority persons in the community sample were too small (see Table 1) for a similar analysis. Within the health center group, mammography use in the past year was higher among those with health insurance coverage than among those without it: 32.9% for those with commercial insurance (with or without other coverage), 23.9% for those with Medicare and/or Medicaid only, and 22.6% for those without insurance ($\chi^2 = 7.658$, $df = 2$, $P = .002$). Among women

TABLE 1—Comparison of Characteristics of Community and Health Center Samples

Characteristic	Community (n = 404) %	Health Center (n = 795) %	Statistical Tests
Age group			$\chi^2 = 32.573$, $df = 4$, $P < .001$
50–54	21.5	13.3	
55–59	28.0	19.5	
60–64	19.3	23.4	
65–69	16.3	23.1	
70–75	14.9	20.6	
Annual family income			$\chi^2 = 235.972$, $df = 4$, $P < .001$
<$15 000	19.3	52.8	
$15 000–24 999	20.3	20.3	
$25 000–34 999	16.3	7.4	
≥$35 000	31.9	4.8	
Unknown	12.1	14.7	
Educational level			$\chi^2 = 76.468$,[a] $df = 3$, $P < .001$
<12 grade	20.1	43.0	
High school grad	50.3	38.6	
Post high school	17.3	14.1	
≥College grad	12.1	4.0	
Unknown	0.3	0.3	
Race/ethnicity			$\chi^2 = 68.678$,[a] $df = 3$, $P < .001$
White/non-Hispanic	91.3	71.5	
Black/non-Hispanic	2.0	11.5	
Hispanic	4.7	15.2	
Other	1.0	1.5	
Unknown	1.0	0.4	

Note. *df* = degree of freedom.
[a] Chi-square calculation excludes unknowns.

TABLE 2—Strongly Agree with Statements about Breast Cancer[a]

Strongly agree with statement:	Community (n = 404) %	Health Center (n = 795) %	Statistical Tests
Mammography can find lumps doctors do not always find	27.2	22.0	$\chi^2 = 3.738$, $df = 1$, $P = .053$
There are other ways to treat besides removing the breast	26.5	17.4	$\chi^2 = 13.168$, $df = 1$, $P < .001$
Women live longer if the cancer is found early	59.2	44.9	$\chi^2 = 21.201$, $df = 1$, $P < .001$

Note. *df* = degree of freedom.
[a] Women were asked to rank the strength of their agreement with the above three statements on a scale of "strongly agree," "agree," "disagree," "strongly disagree," or "don't know." "Strongly agree" is compared with all other responses.

TABLE 3—Concerns about Mammography[a]

Type of Concern	Community (n = 404) %	Health Center (n = 795) %	Statistical Tests
Results	88.9	82.8	$\chi^2 = 7.261$, $df = 1$, $P = .007$
Radiation	73.5	71.1	$\chi^2 = 0.677$, $df = 1$, $P = .411$
Cost	49.0	64.0	$\chi^2 = 24.344$, $df = 1$, $P < .001$
Pain	45.5	47.4	$\chi^2 = 0.307$, $df = 1$, $P = .579$
Embarrassment	34.2	37.5	$\chi^2 = 1.141$, $df = 1$, $P = .285$

Note. *df* = degree of freedom.
[a] Includes any level of concern (i.e., "a little," "somewhat," or "very"). "Any level of concern" is compared with all other responses.

TABLE 4—Awareness and Use of Breast Cancer Screening

Awareness and Use	Community (n = 404) %	Health Center (n = 795) %	Statistical Tests
Ever heard of mammogram	95.3	91.2	$\chi^2 = 5.976$, $df = 1$, $P = .015$
Ever had mammogram	49.3	49.9	$\chi^2 = 0.026$, $df = 1$, $P = .872$
Had mammogram in past year	26.2	29.2	$\chi^2 = 1.007$, $df = 1$, $P = .316$
Ever had clinical breast exam	96.5	95.6	$\chi^2 = 0.385$, $df = 1$, $P = .535$
Had breast exam in past year	57.2	63.8	$\chi^2 = 4.649$, $df = 1$, $P = .031$

Note. df = degree of freedom.

TABLE 5—Mammography Use in Past Year, by Demographic Characteristics

Demographic Characteristics	Community n/total n	Community %	Heath Center n/total n	Heath Center %
Income level				
<$15 000	11/78	14.1	129/420	30.7
$15 000–24 999	19/82	23.2	36/161	22.4
$25 000–34 999	16/66	24.2	16/59	27.1
≥$35 000	46/129	35.7	10/38	26.3
Statistical tests	$\chi^2 = 12.466$, $df = 3$, $P = .006$		$\chi^2 = 4.126$, $df = 3$, $P = .248$	
Education: highest grade completed				
<12 grade	14/82	17.1	94/344	27.3
High school grad	51/203	25.1	95/307	30.9
Post high school	17/70	24.3	30/112	26.8
≥College grad	24/49	49.0	13/32	40.6
Statistical tests	$\chi^2 = 16.921$, $df = 3$, $P = .001$		$\chi^2 = 3.374$, $df = 3$, $P = .337$	
Age group				
50–54	30/8	34.5	29/106	27.4
55–59	33/113	29.2	48/155	31.0
60–64	15/78	19.2	48/186	25.8
65–69	17/66	25.8	61/184	33.2
70–75	11/60	18.3	46/164	28.1
Statistical tests	$\chi^2 = 7.493$, $df = 4$, $P = .112$		$\chi^2 = 2.941$, $df = 4$, $P = .568$	

Note. df = degree of freedom.

TABLE 6—Reasons for Never Having Had a Mammogram, among Women Who Never Had a Mammogram[a]

Reason	Community (n = 204) %	Health Center (n = 398) %	Statistical Tests
Not necessary/no problem	37.3	45.0	$\chi^2 = 2.983$, $df = 1$, $P = .084$
Doctor never recommended	23.0	26.4	$\chi^2 = 0.631$, $df = 1$, $P = .427$
Someone recommended against it	1.0	0.5	$\chi^2 = 0.023$, $df = 1$, $P = .878$
Never thought of it	10.8	8.5	$\chi^2 = 0.560$, $df = 1$, $P = .454$
Cost of mammogram	3.4	7.3	$\chi^2 = 2.912$, $df = 1$, $P = .088$
Concern about radiation	3.4	2.8	$\chi^2 = 0.041$, $df = 1$, $P = .840$
Fear of pain	2.0	1.3	$\chi^2 = 0.102$, $df = 1$, $P = .749$
Other	30.4	22.1	$\chi^2 = 4.512$, $df = 1$, $P = .034$
No reason given	12.2	9.8	$\chi^2 = 0.586$, $df = 1$, $P = .444$

Note. df = degree of freedom.
[a] Includes respondents who, when asked if they had ever had a mammogram, replied "no" or "don't know" or did not answer. More than one reason could be given by respondents.

who had a household income under $15 000, proportionally more than twice as many health center vs community respondents had a mammogram in the past year ($\chi^2 = 8.179$, $df = 1$, $P = .004$). Variations with demographic characteristics also applied to reports of ever having had a mammogram among the community group but not among the health center group. About half (or more) of the health center respondents had ever had a mammogram within each income, age, and race/ethnic subgroup, whereas among the community respondents this proportion varied significantly with income ($\chi^2 = 15.502$, $df = 3$, $P = .001$) and education ($\chi^2 = 20.240$, $df = 3$, $P < .001$).

Among women who had never had a mammogram, the ranking of reasons was similar for both groups of women (Table 6), the most common reason—that it was not necessary or they had no problem—was followed by lack of a doctor's recommendation. More health center women cited cost of the mammogram as a deterrent to mammography, but none of the differences in specific reasons achieved statistical significance.

Discussion

Despite lower levels of knowledge about mammography and the lifetime risk of breast cancer for women, less favorable attitudes regarding the value of early detection and mammography, and sociodemographic characteristics more typical of a disadvantaged population, health center registrants who visited a center in the year prior to the survey used breast cancer screening services at levels comparable to those of a random sample of women in the same community.

In a 1988 study of predominantly Hispanic users of a public health center in Massachusetts, where 59% of the women had ever had a mammogram (compared with 50% for all respondents in our study) and mammography use was similar across age, education, and employment subgroups, there was no comparison group from the general population.[8] In our study, income and educational level also did not significantly influence mammography use within the health center population although they were significant factors within the community; this suggests that services were available more equitably to all subgroups using the health centers. Race/ethnicity and age also did not significantly influence mammography use within the health center group.

The fact that 15% of community women vs 3% of health center women in the lowest household income category (under $15 000) indicated having no regular source of care did not account for the large difference in mammography use between the two samples of women within

this income subgroup. None of the 12 community women with household incomes under $15 000 and no regular source of care had a mammogram in the previous year, compared with 14% (2/14) of health center women in the same subgroup. Among those lowest-income women *with* a regular source of care, however, 31% (127/406) of health center women vs only 17% (11/66) of community women ($P = .02$) had a mammogram in the past year. Thus, having a health center as a source of care was critical.

Both health center and community women reported lower mammography use than is recommended in national guidelines, and they cited concerns that would deter them from obtaining such screening. Inadequate use of screening services for breast cancer reflects both patient and physician factors. Our findings suggest that physicians and others should inform women that the purpose of screening is to detect disease early, before there are symptoms, in order to reduce the high frequency of reports of never having had a mammogram because screening is not necessary or there were no "problems." Another common reason for never having had a mammogram was that the doctor did not recommend it. Research has indicated the importance of physician recommendation for patient compliance with mammography screening.[9–12] But although surveys of primary care physicians have shown some improvement in mammography referral rates since the mid 1980s, there is still inadequate physician compliance with national recommendations for mammography screening.[13–15] Finally, women's concerns about the results of mammography (Table 3) could possibly be alleviated by highlighting the benefits of screening, such as the high curability of breast cancer if detected early and the possibility of lesser surgery.

Use of mammography has been shown to decline with decreasing income in other populations.[4] As expected, cost of mammography was a greater concern among health center than among community respondents. Yet we have shown no difference in mammography use between predominantly low-income health center users and the general population, and significantly higher use within low-income health center users. These findings may be explained, in part, by differences in health care use factors—for example, all of the health center sample had visited the health center in the past year, which often included physician contact, and almost all of them considered the health center to be their regular source of care. Other features of the health center, however, may also have been important and may require examination—for example, physician behavior, such as higher referral rates of low-income women for mammography by health center providers than by other physicians.

Recent legislation that provides some coverage for screening mammography should help increase mammography use in both the community and the health centers. This includes a New York State law effective January 1989 requiring private insurance companies to provide some coverage for screening mammography, and provisions in the Medicare legislation effective in 1991 for Medicare recipients to receive coverage for a screening mammogram every other year.

The results of this study suggest an important role for health centers in accomplishing preventive medicine goals for early detection of breast cancer among the socioeconomically disadvantaged. They also show the potential for achieving equity in behavioral objectives despite lower educational levels and lower levels of understanding about screening, and they provide an example of the importance of community health centers in improving primary and preventive health care in the population served.[16] □

Acknowledgments

This research was supported by grant 1R01 CA 4503401 from the National Cancer Institute. This paper was presented in part at the annual meeting of the American Public Health Association in October 1989.

The authors thank David Kovenock, Ph.D., of Northeast Research (Orono, Me) for assistance in revising the telephone surveys, which were conducted by contract with Northeast Research; and Leijin Chao for assistance in computer data analysis. We are indebted to Drs. David Harris, William Steibel, Judith Feldman, and Clare Bradley of the Suffolk County (NY) Department of Health Services for providing access to the health centers and their registrants, and also to the community advisory boards of the five health centers for approving our survey of the health center population.

References

1. Vanchieri C. Medical groups' message to women: if 40 or older get regular mammograms. *J Natl Cancer Inst.* 1989;81:1126–1128.
2. US Preventive Services Task Force. *Guide to Clinical Preventive Services: An Assessment of the Effectiveness of 169 Interventions.* Baltimore, Md: Williams and Wilkins; 1989;chap 6:39–62.
3. US Dept of Health and Human Services. *Healthy people 2000: National health promotion and disease prevention objectives.* Washington DC: USDHHS; 1990.
4. National Cancer Institute Breast Cancer Consortium. Screening mammography—a missed clinical opportunity? Results of the NCI Breast Cancer Screening Consortium and National Health Interview survey studies. *JAMA.* 1990;284:54–59.
5. Freeman HP. Cancer and the socioeconomically disadvantaged. *CA.* 1989;39:266–288.
6. Lane DS, Polednak AP, Burg MA. Measuring the impact of varied interventions on community-wide breast cancer screening. In: Anderson PN, Engstrom B, Mortonson LE, eds. *Advances in Cancer Control VI.* New York, NY: Alan Liss Inc; 1989:103–114.
7. Polednak AP, Lane DS, Burg MA. Mail versus telephone surveys on mammography utilization among women 50–75 years old. *Med Care.* 1991;29:243–250.
8. Zapka JG, Stoddard A, Barth R, et al. Breast cancer screening utilization by Latina community health center clients. *Health Educ Res.* 1989;4:461–468.
9. Lane DS. Compliance with referrals from a cancer-screening project. *J Fam Pract.* 1983;17:811–817.
10. Burg MA, Lane DS, Polednak AP. Age group differences in the use of breast cancer screening tests: the effect of health care utilization and socioeconomic variables. *J Aging Health.* 1990;2:514–530.
11. Fox SA, Murata PJ, Stein JA. The impact of physician compliance on screening mammography for older women. *Arch Intern Med.* 1990;151:50–56.
12. Zapka JG, Stoddard A, Costanza M, Greene H. Breast cancer screening by mammography: utilization and associated factors. *Am J Public Health.* 1989;79:1499–1502.
13. Lane DS, Burg MA. Breast cancer screening: changing physician practices and specialty variation. *NY State J Med.* 1990;90:288–292.
14. American Cancer Society. A survey of physicians' attitudes and practices in early detection. *CA.* 1985;35:197–213.
15. American Cancer Society. 1989 survey of physicians' attitudes and practices in early cancer detection. *CA.* 1990;40:77–101.
16. Sardell A. *The US Experiment in Social Medicine: The Community Health Center Program, 1965–1986.* Pittsburgh, Pa: University of Pittsburgh Press; 1988.

Geographic and Socioeconomic Variation in the Onset of Decline of Coronary Heart Disease Mortality in White Women

Steve Wing, PhD, Elizabeth Barnett, MSPH, Michele Casper, PhD, and H. A. Tyroler, MD

ABSTRACT

Background. Regional, metropolitan, and socioeconomic factors related to the onset of decline of coronary heart disease (CHD) mortality among White women are reported. Such studies are important for planning population-level interventions.

Methods. Mortality data for 1962 to 1978 were used, to estimate the year of onset of decline. Ecological analyses of socioeconomic data from the US census were used to emphasize structural and organizational aspects of changes in disease, rather than as a substitute for an individual-level design.

Results. Onset of decline of CHD mortality among White women was estimated to have occurred by 1962 in 53% of 507 state economic areas (SEAs), ranging from 79% in the Northeast to 39% in the South. Metropolitan areas experienced earlier onset of decline than did nonmetropolitan areas. Average income, education, and occupational levels were highest in early onset areas and declined across onset categories.

Conclusions. The results provide additional evidence for previously observed geographic and social patterns of CHD decline. Emphasis on structural economic factors determining the shape of the CHD epidemic curve does not detract from the medical importance of risk factors, but underscores the importance of community development to public health improvements. The results are consistent with the idea that the course of the CHD epidemic in the United States has been strongly influenced by socioeconomic development. (*Am J Public Health.* 1992;82:204–209)

Introduction

Early literature suggested that the decline of coronary heart disease (CHD) mortality began around the mid-1960s in all sociodemographic groups in the United States.[1–3] However, declining CHD mortality had been noted in California and New York a decade before the national decline began,[4,5] and a later study demonstrated evidence of substantial variation in the onset of decline among the 50 states for White women and men.[6] Our previous work has shown that geographic variation in the onset of CHD decline among White men in the United States was related to the metropolitan and socioeconomic characteristics of areas.[7,8] Regional, metropolitan, and socioeconomic characteristics related to the onset of decline of CHD mortality among White women aged 35 to 74 are reported here.

In this study we use populations as units of analysis; this design is typically called "aggregate" or "ecological." Ecological studies are generally considered to be inferior substitutes for more powerful individual designs, ideally a randomized experiment, in which information on both exposure and outcome is known for each person.[9] Although ecological studies are clearly inferior when the purpose of research is to isolate "independent" risk factor–disease associations, they are essential for addressing population characteristics such as epidemic cycles or economic relations. The ecological design is chosen here not as a substitute for a "superior" individual study, but as a tool for investigating population-level determinants of disease that cannot be addressed with individual-level study designs. An understanding of these determinants can help to delineate the public health impact of policies in areas such as economic development, housing, or agriculture, and their potential for reducing mass disease.

Methods

Individual records of underlying cause of death for all US residents for the years 1962 through 1978 were used. Because the definition of CHD was changed between the seventh (1962 to 1967) and eighth (1968 to 1978) revisions of the International Classification of Diseases (ICD), a group of seventh-revision codes believed to show good comparability with the eighth-revision definition[10] were chosen: 420, 422.1, 440, 441, and 443. The standard eighth-revision definition of 410–413 was used.

CHD deaths for White women aged 35 to 74 were grouped in 5-year age categories within each state economic area (SEA) in the continental United States. SEAs are single metropolitan counties or groups of nonmetropolitan counties. Annual SEA age-specific population counts were derived by linear intercensal interpolation for the 1960s and were available from the Census Bureau for the 1970s.

At the time of this study the authors were with the Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill. Elizabeth Barnett is now with the Division of Maternal and Child Health, North Carolina Department of Environment, Health and Natural Resources, Raleigh, NC. Michele Casper is now with the Centers for Disease Control, Atlanta, Ga.

Requests for reprints should be addressed to Steve Wing, PhD, Department of Epidemiology, School of Public Health, CB# 7400, McGavran-Greenberg Building, University of North Carolina, Chapel Hill, NC 27599-7400.

This paper was submitted to the journal November 9, 1990, and accepted with revisions July 17, 1991.

Age-adjusted rates for each SEA were computed by the direct method, using the total 1970 US population as the standard. More detailed information about data sources has been published previously.[11]

Quadratic regression was used to model the annual age-adjusted rates for each SEA.[6-8,12] In this model the adjusted rate is considered to be a function of an intercept, a slope due to year and a slope due to year squared. SEAs with declining rates over the entire period were considered to have experienced onset of decline in 1962 or earlier. Other SEAs were considered to have experienced onset of decline in the year of the maximum estimated rate from the quadratic regression.

The distribution of the timing of the onset of decline of CHD mortality was examined by geographic region and metropolitan status. The four standard US census region definitions were used: Northeast, South, North Central, and West.[7] SEAs were classified as metropolitan (n = 206) if the component counties were part of a standard metropolitan statistical area with a central city of more than 50 000 or a total population of 100 000 or more. Other SEAs were classified as nonmetropolitan (n = 301). Average levels of socioeconomic characteristics from the 1960 and 1970 US censuses were used to characterize each SEA. On the basis of previous work,[8] we considered three measures: percentage of families with annual income greater than $10 000, percentage of the labor force employed in white-collar occupations, and percentage of the population aged 25 and above that had completed high school or more. These measures reflect many aspects of individuals and their communities, and are considered here primarily as indicators of socioeconomic development and community resources. For example, high levels of white-collar employment reflect an area's economic base in service, government, and financial industries rather than in heavy manufacturing, agricultural, or extractive industries. White-collar employment, as an indicator of occupational structure, is as much a function of the kinds of jobs provided by a community as it is of the characteristics of the workers. The ecological use of these measures stresses not only the implications of white-collar employment for white-collar workers, but the effect of the occupational structure on health-relevant living conditions—including diet, recreation, housing, working conditions, and medical care—for the entire population.

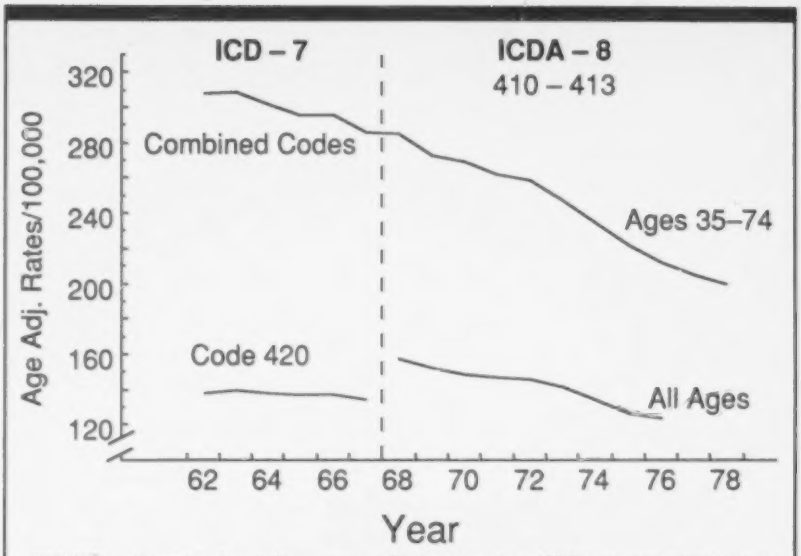


FIGURE 1—Coronary heart disease mortality trends among White women in the United States, 1962–1978: Comparison of coding schemes across ICD revisions 7 and 8.

Results

The comparability of CHD rates when two seventh-revision definitions and the standard eighth-revision definition were used is shown in Figure 1. The bottom line shows published direct-adjusted rates for all ages for ICD-7 code 420.[10] The discontinuity evident in that trend is eliminated by the recoding of seventh-revision data (codes 420, 422.1, 440, 441, and 443) shown in the top line for ages 35 to 74. Similar graphs of the recoded data for individual SEAs showed no evidence of discontinuities due to geographic variation in the comparability of the CHD definitions.

Quadratic regression estimates of the CHD mortality trend within each SEA for the years 1962 to 1978 showed that in 271 (53%) of the 507 SEAs, in which 70% of White women aged 35 to 74 resided, the highest estimated CHD rate occurred in 1962. These SEAs were classified as experiencing onset of decline in 1962 or earlier. Figure 2 shows the cumulative percentage distributions of onset of decline by metropolitan status. CHD rates in 59% of metropolitan SEAs, compared with 50% of nonmetropolitan SEAs, were estimated to have begun to decline by 1962. The earlier timing of the onset of decline in metropolitan areas was maintained through most of the distribution.

Figure 3 shows cumulative percentage distributions of onset of decline by region. The Northeast region contained the largest proportion of SEAs in onset of decline by 1962, and it remained ahead of other regions throughout the period of onset of decline. The North Central and West regions showed almost identical distributions of onset of decline; the South was later than other regions in the onset of decline. Metropolitan differentials seen for the United States were observed within each region. The more urban composition of the Northeast and the more rural composition of the South, however, did not account for the differences between these regions and the US average: 82% of metropolitan and 74% of nonmetropolitan SEAs in the Northeast were in decline by 1962, compared with 42% and 38%, respectively, in the South. Onset of decline was estimated to have occurred after 1968 in 40 SEAs, of which 35 were in the South. The Appendix gives percentages of SEAs experiencing onset of decline of CHD mortality in different time periods by metropolitan status and region.

Figure 4 shows average levels of 1970 socioeconomic indicators among SEAs grouped into four categories of onset of decline: 1962 or earlier, 1963 to 1965, 1966 to 1968, and 1969 or later. SEAs in the earliest onset-of-decline category show the highest levels of each of the three measures (percentage of families with annual income greater than $10 000, proportion of the labor force in white-collar jobs, and proportion of the population aged 25 and above with a high school diploma). SEAs in later onset-of-decline categories show progressively lower proportions of the population

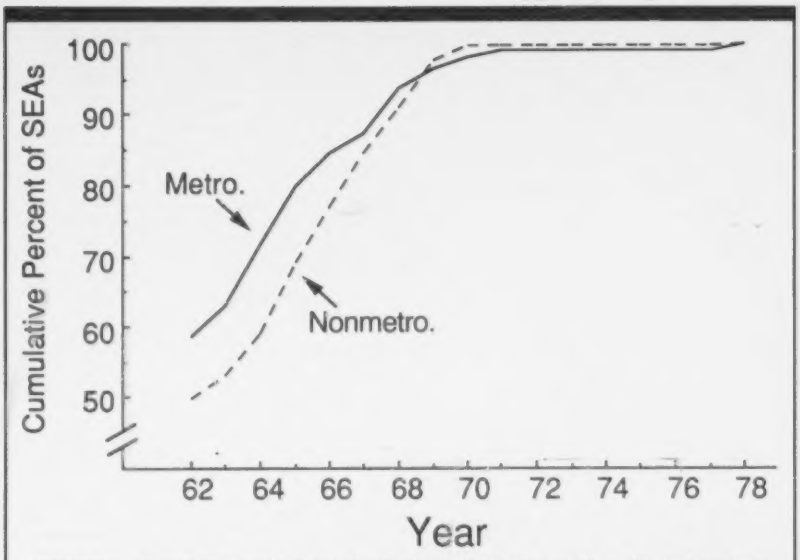


FIGURE 2—Cumulative percentage distributions of onset of decline of coronary heart disease mortality among White women aged 35–74, by metropolitan status.

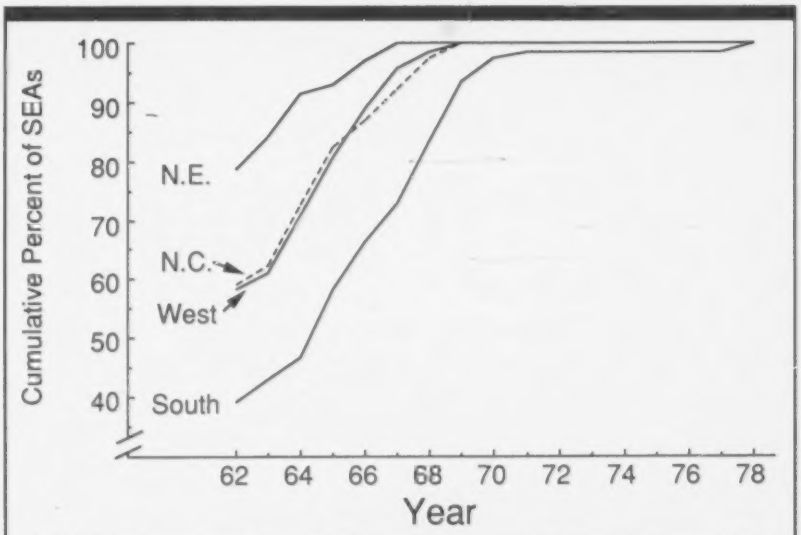


FIGURE 3—Cumulative percentage distributions of onset of decline of coronary heart disease mortality among White women aged 35–74, by region.

with high levels of these indicators. An examination of the average socioeconomic characteristics of SEAs by onset of decline separately within region and metropolitan status categories (data not presented) showed differences in the same direction but of smaller magnitude. Analyses using 1960 census indicators of socioeconomic conditions showed similar results.

The estimated time period of onset of decline of each of the 507 SEAs of the continental United States is shown on the map (Figure 5). Variability in onset of decline is evident in all regions and in most states. Spatial concentrations of early onset of decline are evident in the Northeast and West, whereas concentrations of late-onset areas appear in the South, especially in the area from Oklahoma through the Mississippi Delta to Georgia and in eastern Tennessee and Kentucky.

Discussion

In general, areas in the United States experiencing earlier onset of decline of CHD mortality among White women showed higher levels of socioeconomic indicators than areas experiencing later onset of decline. These early-onset areas tended to be metropolitan areas and areas outside the South. This pattern of onset of decline is similar to the pattern reported earlier for White men in the United States[7,8] and also to the pattern reported for White men in the Netherlands, where data for women were not analyzed.[12] The current results provide additional quantitative evidence for the previously observed geographic patterns.

Results of such ecological studies are often interpreted as reflecting differentials occurring among individuals; for example, it may be argued that places have earlier onset of decline because there are greater proportions of individuals of higher socioeconomic status there. Some evidence supports this interpretation for CHD mortality among men[13,14] and for changes in smoking behavior,[15] an important risk factor for CHD. However, ecological studies remain a poor method of isolating individual risk factor–disease associations, owing to the many potential ambiguities of that interpretation.[9]

The information used to evaluate the plausibility of findings about risk factor–disease relationships is usually derived primarily from genetic, behavioral, physiological, and clinical disciplines. These epidemiological perspectives emphasize aspects of individuals rather than the organization of individuals in groups, and they focus on the roles of agents (physical, biological, or psychosocial) in the production of disease in individuals. The potential strength of the ecological approach, on the other hand, is in understanding the production of disease in populations viewed not merely as numbers of individuals, but as organized groups with structures and relationships that create health and disease. In order to realize this strength, epidemiological information on risk factors must be combined with perspectives on populations as organized systems. These perspectives are better developed in areas such as ecology, evolutionary biology, history, sociology, and political economy than in molecular and clinical disciplines. An ecological approach to epidemiology can use these perspectives to provide an alternative account of the causes of disease in populations.

Ecological studies are crucial to a public health epidemiology (as it may now be distinguished from a clinical epidemiology) because they suggest mechanisms

of disease production and possibilities for intervention that involve human populations as organized systems.[16,17] In the case of the study of geographic variation in the onset of decline of CHD mortality, we recognize that areas experience historical economic and social development with profound implications for industrial structure, local wages, educational resources, food supplies, recreational opportunities, and ability to attract medical personnel and facilities. Studies of the mechanisms through which these processes may affect population distributions of risk factors and disease trends are part of developing the ecological view.

In the United States, the Northeast has traditionally been a rich economic area in relation to the poorer South, which has provided raw materials and cheap labor to capital centers of the Northeast.[18,19] Despite some evidence of regional convergence of incomes during the 1960 to 1976 time period, especially for advantaged workers, a substantial South–non-South gap was still present for blue-collar and less educated workers.[20] Additionally, a metropolitan–nonmetropolitan gap in industrial base, occupational structure, wages, and living conditions has developed within the South.[21] Interestingly, the metropolitan differential in the onset of decline of CHD in White women in the South appears to be smaller than the metropolitan differential in other regions (Appendix); however, the growing inequality of CHD mortality in White men and women between areas classified on the basis of occupational structure is clearer in the South than in other regions of the country.[22] As we have noted previously,[7] factors may affect the onset of decline and the rate of decline in different ways.

The delayed onset of decline of CHD mortality in the South compared with the rest of the United States (Figure 3) is consistent with the South's historical underdevelopment. This "regional colonialism"[18] has been implicated in epidemic patterns of hookworm and pellagra in the early 20th century,[23,24] and suggests continuity of structural determinants of health during the transition from "old" to "new" diseases as the major public health problems. Such historical continuity of disease causation in populations implies an alternative understanding of epidemiological transition, which has traditionally been seen as involving a parallel change in the *causes* of health problems and the major diseases of a particular period.[25] With respect to understanding and reducing poor

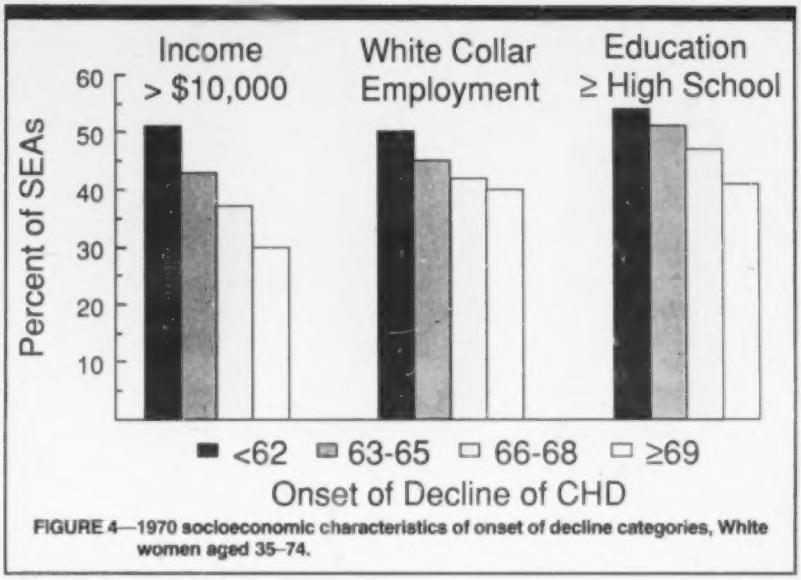


FIGURE 4—1970 socioeconomic characteristics of onset of decline categories, White women aged 35–74.

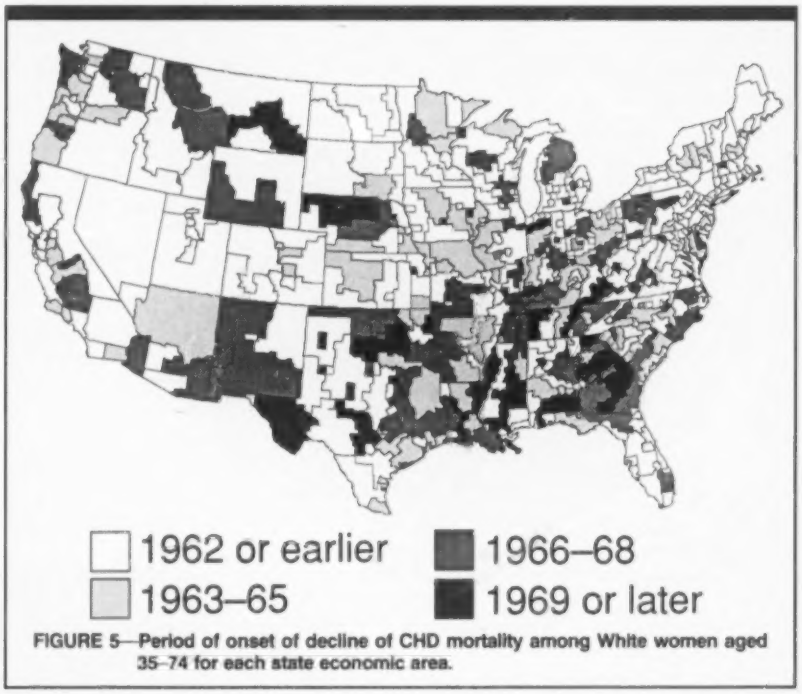


FIGURE 5—Period of onset of decline of CHD mortality among White women aged 35–74 for each state economic area.

health conditions in the South, the major diseases and disease agents may change while the population-level relations that determine their distribution are maintained. An analogous situation could be described in the experience of epidemiologic transition in some underdeveloped countries.[26]

When population relationships of development and underdevelopment are stable through epidemiological transition, the association between population characteristics and disease can change dramatically. For example, areas that experienced earlier onset of decline of CHD mortality in the period covered in this study showed a more rapid increase of cardiovascular mortality during the 1930s.[27] The epidemic of CHD accompanied rising levels of affluence that brought more atherogenic diets, physical inactivity, and increased cigarette smoking. Ironically, this relative affluence resulted from hard manual labor and lower consumption

by underdeveloped populations in the United States and abroad, factors that simultaneously limited the CHD epidemic in these populations but maintained high levels of other diseases. Now, while risk factor reduction is taking place in some more affluent populations, smaller reductions or increases are simultaneously occurring in less developed areas.

The public health potential of these ecological interpretations rests on the assumption that community resources and public policies condition options and choices of behavior in diet, smoking, physical activity, and medical care utilization (especially in a country without a national medical care system).[28] A number of theorists have suggested specific forms of economic development that should be conducive to health, such as emphasis on local ownership to prevent the outflow of profits and to reduce incentives for industries to relocate to take advantage of cheaper labor or materials.[29] Although economic conditions related to basic living conditions may have the greatest potential for improving public health, such broad issues have political dimensions that limit their potential as direct public health interventions. Those changes will occur only as a function of a more broadly conceived social agenda.

Other ecological aspects of risk factor and disease distributions, however, may offer more realistic possibilities as public health interventions, because they involve medically recognized disease agents. For example, cigarette smoking and dietary fat are relevant to the epidemic pattern of CHD. It has been well documented that declines in smoking among higher-status individuals have been followed by more aggressive marketing of tobacco to the poor, racial minorities, and women, both in the United States and abroad.[30,31] The geographic dimension of this marketing process may contribute to the patterns of onset of decline of CHD noted here as well as to growing geographic inequalities in CHD mortality. Empirical analyses documenting geographic variation in trends in tobacco advertising and consumption, if consistent with this interpretation, would support public health policies aimed at controlling tobacco production and marketing, because such interventions would protect against the shifting of consumption from populations in more developed to those in less developed areas.[32] A similar situation could be described for production and marketing of high-fat foods, which are being forsaken, to some degree, by the better educated and are being marketed in less developed areas. This situation extends even to the distribution of free dairy products, such as cheese, in government programs. These ecological mechanisms suggest that interventions in agricultural policy and marketing could be effective environmental controls on the population distribution of heart disease.[33] Because they involve more focused mechanisms than economic interventions targeted directly at community development, they constitute more realistic short-term health policy objectives.

The results presented here suggest that the onset of decline of CHD mortality in White women has been related to historical patterns of economic development in the United States. These analyses represent the first evidence, to our knowledge, of such patterns among women, inasmuch as previous analyses for the United States and the Netherlands excluded women.[7,8,11] Internationally, CHD mortality levels and trends have shown less variability for women than for men,[34] but the onset of decline of CHD mortality in the United States appears to show at least as much variability in women as in men, and a clear pattern of differences between the South and other regions. We use the ecological study design to stress structural mechanisms operating in populations rather than to isolate risk factor–disease associations that could be better studied through an individual-level design. This epidemiological approach focuses on the determinants of the population distribution of disease rather than on the identification of specific risk factors, and draws attention to mechanisms involving the organization and relations of populations in addition to the behaviors, exposures, and pathologies of individuals; the logic is economic and environmental as well as clinical and behavioral. Future empirical work documenting the ecological determinants of the distributions of known risk factors could contribute to the development of preventive measures to create healthier environments. □

Acknowledgments

This work was supported by NIH grant 5-R01-HL42320-01. Ms Barnett was supported by a graduate fellowship from the National Science Foundation.

Some of the material in this paper was presented at the 118th annual meeting of the American Public Health Association in New York City, September 30 to October 4, 1990.

The authors thank Naomar Almeida-Filho for comments on an earlier draft of the paper.

References

1. Havlik RJ, Feinleib M, eds. *Proceedings of the Conference on the Decline of Coronary Heart Disease Mortality.* Washington, DC: NIH; 1979.
2. Cooper R, Stamler J, Dyer A, et al. The decline in mortality from coronary heart disease, USA, 1968–1975. *J Chron Dis.* 1978;31:709–720.
3. Stallones RA. The rise and fall of ischemic heart disease. *Sci Am.* 1980;243:53–59.
4. Borhani NO, Hechter HH. Recent changes in CVR disease mortality in California. *Pub Health Rep.* 1964;79:146–160.
5. Winkelstein W. Some ecological studies of lung cancer and ischemic heart disease mortality: regional variation. *Int J Epidemiol.* 1985;14:39–47.
6. Ragland KE, Selvin S, Merrill DW. The onset of decline in ischemic heart disease mortality in the United States. *Am J Epidemiol.* 1988;127:516–531.
7. Wing S, Hayes C, Heiss G, et al. Geographic variation in the onset of decline of ischemic heart disease mortality in the United States. *Am J Public Health.* 1986;76:1404–1408.
8. Wing S, Casper M, Riggan W, et al. Socioenvironmental characteristics associated with the onset of decline of ischemic heart disease mortality in the United States. *Am J Public Health.* 1988;78:923–926.
9. Morgenstern H. Uses of ecologic analysis in epidemiologic research. *Am J Public Health.* 1982;72:1336–1344.
10. Rosenberg HM, Klebba AJ. Trends in cardiovascular mortality with a focus on ischemic heart disease: United States, 1950–1976. In Havlik RJ, Feinleib M, eds. *Proceedings of the Conference on the Decline of Coronary Heart Disease Mortality.* Washington, DC: NIH; 1979.
11. Wing S, Casper M, Davis W, et al. Trends in the geographic inequality of cardiovascular disease mortality in the United States, 1962–1982. *Soc Sci Med.* 1990;30(3):261–266.
12. Mackenbach JP, Looman CWN, Kunsa AE. Geographic variation in the onset of decline of male ischemic heart disease mortality in the Netherlands. *Am J Public Health.* 1989;79:1621–1627.
13. Pell S, Fayerweather WE. Trends in the incidence of myocardial infarction and in associated mortality and morbidity in a large employed population, 1957–1983. *N Engl J Med.* 1985;312:1005–1011.
14. Feldman JJ, Makuc DM, Kleinman JC, Cornoni-Huntley J. National trends in educational differentials in mortality. *Am J Epidemiol.* 1989;129:919–933.
15. Schoenborn CA, Cohen BH. Trends in smoking, alcohol consumption, and other health practices among U.S. adults, 1977 to 1983. *Advance Data 118,* US National Center for Health Statistics; 1986.
16. Loomis D, Wing S. Is molecular epidemiology a germ theory for the end of the twentieth century? *Int J Epidemiol.* 1980;19:1–3.
17. Illsley R. Comparative review of sources, methodology and knowledge. *Soc Sci Med.* 1990;31:229–236.

18. Perksy J. Regional colonialism and the Southern economy. *Rev Rad Polit Econ.* 1972;5:70–79.
19. Fox K. Uneven regional development in the United States. *Rev Rad Polit Econ.* 1978;10:68–86.
20. Hirschman C, Blankenship K. The North–South earnings gap: changes during the 1960s and 1970s. *Am J Sociology.* 1981;87(2):388–403.
21. Falk WW, Lyson TA. *High Tech, Low Tech, No Tech. Recent Industrial and Occupational Change in the South.* Dworkin AG, ed. State University of New York Press; 1988.
22. Wing S, Casper M, Hayes CG, Dargent-Molina P, Riggan W, Tyroler HA. Changing association between community occupational structure and ischaemic heart disease mortality in the United States. *Lancet.* November 7, 1987;ii:1067–1070.
23. Brown ER. Public health in imperialism: early Rockefeller programs at home and abroad. *Am J Public Health.* 1976;66:897–903.
24. Kunitz SJ. Hookworm and pellagra: exemplary diseases in the New South. *J Health Soc Behav.* 1988;29:139–148.
25. Omran AR. Epidemiology transition in the U.S.: The Health Factor in Population Change. *Population Bull.* 1977;32(2):3–41.
26. Frenk J, Bobadilla JL, Sepulveda J, Cervantes ML. Health transition in the middle-income countries: new challenges for health care. *Health Policy Planning.* 1989;4:29–39.
27. Goves M. Statistical studies of heart disease: IV. Mortality from heart disease (all forms) related to geographic section and size of city. *Pub Health Rep.* 1949;64:439–456.
28. Milio N. Making healthy public policy; developing the science by learning the art: an ecological framework for policy studies. *Health Promotion* 1988;2:263–274.
29. Sclar ED. Community economic structure and individual well-being: a look behind the statistics. *Int J Health Serv.* 1980; 10:563–579.
30. Davis RM. Current trends in cigarette advertising and marketing. *N Engl J Med.* 1987;316:725–732.
31. Barry M. The influence of the U.S. tobacco industry on the health, economy, and environment of developing countries. *N Engl J Med.* 1991;324:917–920.
32. Wing S. Social inequalities in the decline of coronary mortality. *Am J Public Health.* 1988;78:1415–1416.
33. Milio N. Promoting health through structural change: analysis of the origins and implementation of Norway's farm-food-nutrition policy. *Soc Sci Med.* 1981;15A:721–734.
34. Uemura K, Pisa Z. Trends in cardiovascular disease mortality in industrialized countries since 1950. *World Health Stat Q.* 1988;41:155–178.

APPENDIX—Distribution of State Economic Areas (SEAs) According to Estimated Period of Onset of Decline of Coronary Heart Disease Mortality, by Region and Metropolitan Status

Region	*n*	Percentage of SEAs			
		1 1962 or Earlier	2 1963–1965	3 1966–1968	4 1969 or Later
U.S.	507	53.5	20.1	18.5	7.9
Metro	206	58.7	21.4	13.6	6.3
Nonmetro	301	49.8	19.3	21.9	9.0
Northeast	70	78.6	14.3	7.1	0.0
Metro	39	82.1	15.4	2.6	0.0
Nonmetro	31	74.2	12.9	12.9	0.0
North Central	154	59.1	23.4	14.9	2.6
Metro	61	83.9	21.3	13.1	1.6
Nonmetro	93	55.9	24.7	16.1	3.2
South	211	39.3	19.0	25.1	16.6
Metro	79	41.8	21.5	21.5	15.2
Nonmetro	132	37.9	17.4	27.3	17.4
West	72	58.3	22.2	18.1	1.4
Metro	27	63.0	29.6	7.4	0.0
Nonmetro	45	55.6	17.8	24.4	2.2

Note. Thirty-five of the 40 SEAs in category 4 are in the South.

Vaginal Douching among Women of Reproductive Age in the United States: 1988

ABSTRACT

Background. Vaginal douching has been associated with pelvic inflammatory disease (PID) in several epidemiologic studies.

Methods. To determine the extent to which douching is practiced and to describe the population subgroups in which it is most prevalent, we analyzed data from the 1988 National Survey of Family Growth, which is based on a nationally representative sample of 8450 United States women between the ages of 15 and 44 years.

Results. Thirty-seven percent of the sample reported douching; 18% douched at least once a week. The variable most strongly and consistently associated with douching was race: two thirds of Black women, but only one third of White women, reported douching. The practice was least frequent among 15- to 19-year-olds (31%) and most frequent among 20- to 24-year-olds (41%). Douching was more common among women who lived in poverty (50%) than among those who did not (28%). Seventy percent of Black women living in poverty reported douching. Women with less than a high school education were almost four times more likely to report douching as those with 16 or more years of schooling (56% vs 16%). Women with only 1 partner and those with 10 or more partners were less likely to douche than others. Sixteen percent of women who reported douching, compared with 10% of those who did not, also reported a history of PID.

Conclusions. Douching may be a modifiable risk factor for PID, it should be a high priority for future etiologic research. (*Am J Public Health.* 1992;82:210–214)

Sevgi Okten Aral, PhD, William D. Mosher, PhD, and Willard Cates, Jr, MD, MPH

Introduction

Vaginal douching is an ancient practice traditionally used for cleansing the vagina. It is apparently common today. In 1984, 32% of women over the age of 18 reported having douched within the previous week.[1] Almost 70% of sexually active adolescent women who learned about douching from their mothers practiced it.[2] However, douching may be dangerous. The practice has been associated with pelvic inflammatory disease (PID)[1,3–5] and ectopic pregnancy.[6,7] Douching remained a statistically significant risk factor for these conditions, even after the effects of demographic factors, microbiologic factors, sexual history, and other behavioral factors were controlled. Mechanisms by which douching might facilitate PID include altering the vaginal environment toward more pathogenic organisms and flushing vaginal and cervical pathogens into the uterus and fallopian tubes.[8]

If douching is in fact a risk factor for PID, a health education campaign against the practice may help to reduce PID and its sequelae. The impact of such preventive interventions depends on the prevalence of douching in the population; the greatest effect will be observed in those subgroups in which the practice of douching is most widespread. The prevalence of douching varies both geographically and demographically throughout the United States.[9,10] In selected studies, the practice has been reported more frequently by Black women than by others.[1–3]

In this article, we report our findings on the prevalence of vaginal douching among a nationally representative sample of American women of reproductive age. We discuss the distribution of douching practices by demographic (age, race, marital status, education, and poverty level), ecologic (geographic region and urban residence), behavioral (age at first intercourse, number of lifetime sexual partners), and medical (past history of sexually transmitted disease [STD] and PID) characteristics. We also present data on the frequency of douching and its timing in relation to sexual intercourse.

Methods

This analysis is based on data collected by the National Survey of Family Growth (NSFG), Cycle IV, conducted in 1988 by the National Center for Health Statistics. The NSFG, a periodic survey of women of reproductive age, provides information on contraception, infertility, reproductive health, and prenatal care. The Cycle IV survey was conducted with a multistage probability sample of 8450 women of reproductive age. The sample included 5354 White women, 2771 Black women, and 325 women of other races. The data were collected through structured personal interviews. Response rates were high overall (82.5%), and were similar for Black (82.2%) and White (82.6%) women.

The percentages reported in this article are weighted national estimates. The

The authors are with the Centers for Disease Control, Division of STD/HIV Prevention, Atlanta, Ga, and the National Center for Health Statistics and National Survey of Family Growth.

Requests for reprints should be addressed to Sevgi Okten Aral, PhD, Division of STD/HIV Prevention (E-44), Center for Prevention Services, Centers for Disease Control, Atlanta, GA 30333.

This paper was submitted to the journal December 11, 1990, and accepted with revisions June 10, 1991.

TABLE 1—Percentage of Women Who Douche Regularly, by Demographic and Ecologic Characteristics

	All Races	White	Black
Age, y			
15–19	31.0*	25.4*	53.5**
20–24	41.1	35.7	63.1
25–29	37.6	32.9	67.6
30–34	36.0	31.5	64.8
35–39	35.1	30.2	70.2
40–44	37.0	33.8	65.8
Total	36.7	32.0	66.5
Marital status			
Currently married	32.7*	30.3	63.5
Never married	38.4*	29.1*	67.8
Formerly married	48.4	44.2	68.4
Years of schooling[a]			
0–11	56.4*	51.5*	75.5
12	45.5*	42.1*	69.5**
13–15	31.4*	26.1*	66.0
16+	16.7	13.6	53.9
Poverty level			
Below poverty	50.1***	40.1	69.7
100–199	42.2**	37.6	68.5
200–399	37.1*	33.5*	65.4
400+	28.2	26.1	60.1
Place of residence			
Central city	41.2*	30.8	66.8
Suburb	33.6*	30.7***	65.2
Nonmetropolitan	39.5	36.2	68.3
Geographic region			
All Regions	36.7	32.0	66.5
West	28.1	26.1	57.9
Northeast	31.4	28.1	62.1
Midwest	32.1*	27.4*	68.2
South	48.0	42.3	68.3

[a]These percentages exclude women < 22 years of age.
*The difference between this percentage and the percentage below is significant at $P < .001$.
**The difference between this percentage and the percentage below is significant at $P < .05$.
***The difference between this percentage and the percentage below is significant at $P < .01$.

weight for each respondent represents her probability of selection for the survey, adjusted for nonresponse from women with similar characteristics and adjusted to agree with United States Bureau of the Census estimates by age, race, marital status, and parity. For the first time since 1976, the 1988 NSFG included questions on vaginal douching. Respondents who had ever had intercourse were asked if they douched regularly; if they douched after sexual intercourse, at other times, or both; and how often they douched. We excluded from our analysis 16 women for whom we had no data on douching. Thus, our analysis is based on 7603 women who had ever had intercourse—4786 White, 2554 Black, and 263 of other races. The data represent national estimates of the percentage of sexually experienced women who douche regularly.

We used normal deviate tests (Z tests) to test the statistical significance of differences between proportions. We estimated sampling errors using the balanced half-sample replication technique.[11] This technique takes into account the sample design on which the weighted percentages are based. It does *not* assume simple random sampling. To determine which variables were independently associated with douching when the effects of other variables were controlled, we used multiple logistic regression.

We conducted the multiple logistic regression analysis first for all women, using race as a predictor variable. Since race was such a powerful predictor, we repeated the logistic regression analysis separately for Black and White women. We ran each logistic regression equation two ways: first with unweighted data, and second with the weights to correct for the differential sampling rates and approximate number of sample cases. The results were similar using both approaches. We present the results of the weighted analysis.

Results

Overall, in 1988, over one third (37%) of women of reproductive age reported douching (see Table 1). Race was strongly and consistently associated with douching practices. Two thirds of Black women, but only one third of White women, reported douching. The prevalence of douching by women of either race was lowest for teenagers. However, after age 19, the proportion douching increased by 10 percentage points for both Black and White women and remained quite stable thereafter. More formerly married women than currently married women reported douching; the never-married fell in the middle.

Education was inversely related to douching, and the effect of education was greater among White than among Black women. White women with less than a high school education were almost four times more likely to douche than White college graduates, whereas Black women with less than a high school education were only 40% more likely to douche than Black college graduates. Poverty was also inversely related to douching. As with education, the effect of poverty was greater among Whites than among Blacks. White women living in poverty were 54% more likely to douche than those in the highest income group, while Black women living in poverty were only 16% more likely to do so.

Slightly more of the women residing in nonmetropolitan areas douched than did those living in metropolitan areas. White women living in nonmetropolitan areas were more likely to douche than were White women living in metropolitan areas. Differences among Blacks were not statistically significant. The lowest percentage of women reporting douching lived in the western United States, and the highest percentage lived in the South. About 42% of White women living in the South reported douching regularly, compared with 26% to 28% of those who live in other regions ($P < .001$). Differences were smaller among Blacks. The interregional variability in vaginal douching practices was greater for Whites (range = 16 percentage points) than for Blacks (range = 10 percentage points).

Vaginal douching was more common among women who engaged in sexual intercourse earlier in life (see Table 2). Compared with members of the same race who had their first sexual experience after age 19, White women who began having sexual intercourse before age 15 were 1.6 times more likely, and Black women who began having sexual intercourse before age 15 were 1.2 times more likely, to douche regularly.

The relationship between douching and number of lifetime sex partners was bell-shaped. Women who had had only 1 partner were the least likely to douche, and women with 2 to 9 partners were the most likely to douche; those who reported 10 or more partners fell in the middle. This pattern was consistent for both Blacks and Whites.

Past history of STD and past history of PID were associated with douching patterns in opposite ways. Women who had a history of STD (gonorrhea, chlamydia, herpes infection, or genital warts) were less likely to douche than those who did not report a history of STD. Conversely, women who reported a history of PID were more likely to douche regularly than were women who did not have a history of PID. The patterns of association between douching and past history of PID and STD were consistent for both races.

Among women who reported douching regularly, most (71%) douched infrequently—once a week or less often (see Table 3). Black women who douched tended to douche more frequently than White women. More than one third of Blacks, but only one fourth of Whites who douched did so two or more times per week. Most women who douched reported douching both after intercourse and at other times. Only 5% reported douching only after intercourse. Of the women who douched following sexual intercourse, 36% did so within 30 minutes.

In the multiple logistic regression analysis for all women, race emerged as the single most important predictor. When the effects of age, education, income, age at first sexual intercourse, and number of lifetime partners were controlled, the analysis showed that almost four times as many Black women as White women douched regularly. Education, age, and age at first intercourse also had statistically significant effects on douching. Income and number of lifetime sex partners did not have any significant effect on douching when the other demographic and behavioral variables were controlled.

TABLE 2—Percentage of Women Who Douche Regularly, by Behavioral Characteristics

	All Races	White	Black
Age at first intercourse, y			
Under 15	45	36	68
15–17	43*	38*	67
18–19	35*	31*	68**
20+	25	22	58
Number of lifetime sex partners			
1	30*	28*	54*
2–9	41*	35	79
10+	36	32	63
Past history of STD			
None	37	32	67
Gonorrhea, chlamydia, or herpes	35	29	59
Genital warts	29	26	61
Past history of PID			
No	35*	31*	65*
Yes	47	41	72

*The difference between this percentage and the percentage below is significant at $P < .001$.
**The difference between this percentage and the percentage below is significant at $P < .05$.

TABLE 3—Prevalence of Specific Douching Practices

	All Women (%)	White %	Black %
Do not douche	63.3	68.0	33.5
Douche regularly	36.7	32.0	66.5
Frequency[a]			
2 times per week	10.5	8.2	24.2
1 time per week	7.7	7.1	11.9
2–3 times per month	10.7	9.2	21.2
≤ 1 time per month	7.7	7.5	9.3
Timing in monthly cycle[a]			
After intercourse	5.5	4.4	9.0
Other times	45.5	50.4	31.4
Both	49.0	45.2	59.6
Timing after intercourse[b]			
Within 30 minutes	36.4	32.3	42.6
After 30 minutes	63.6	67.7	57.4

Note. All White–Black differences in this table are statistically significant at $P < .001$, except for the difference for frequency of douching ≤ 1 time per month, which is statistically significant at $P < .05$.
[a]These figures represent percentages of women who douche regularly.
[b]These figures represent percentages of women who douche after intercourse.

Multivariate analysis showed that for all races, factors significantly associated with douching were age, race, education, and age at first intercourse. The results for White women were similar to those for all races. When all other variables were controlled, education had the greatest effect on likelihood of douching. Age and age at first intercourse also emerged as independent predictors of douching with statistically significant effects. Number of lifetime sex partners and income had no independent effect when the other variables were controlled. For Black women, only education and age at first intercourse had independent, statistically significant effects when the other variables were controlled. The demographic and behavioral characteristics we examined did not pre-

dict douching as well for Blacks as they did for Whites, as suggested by the smaller chi-square value for Blacks (34.6) than for Whites (327.8).

Discussion

We found a relatively high prevalence of vaginal douching (37%) among women of reproductive age. These are the first nationally representative, systematically collected data on levels of douching. Only one other investigation, a prevalence study sponsored by a manufacturer, had surveyed douching levels in a nationwide sample of women. According to that marketing survey, conducted in 1984, 32% of women had douched during the previous week,[1] a prevalence remarkably consistent with our finding of 37%. Other surveys of convenience samples had found douching prevalences ranging from 43% to 57%.[9,10]

Race was the most important predictor of douching. Prevalence was particularly high among Blacks. Education and age at first intercourse, the two independent predictors of douching among both White and Black women, were more strongly correlated with douching among White than among Black women. We believe the high prevalence of the practice among Blacks did not allow much variability across categories of any of the predictor variables, and made it difficult for any demographic, ecologic, or behavioral variable to emerge as a powerful predictor.

In a recent report, douching emerged as an important risk factor for PID among White women but not among Black women, due to the high prevalence of the practice among Black controls.[3] Further analyses of the NSFG data indicate that douching is independently associated with PID, even when number of sex partners, age, STD history, age at first intercourse, and marital status (among White women, but not among Black women) are controlled.[12] Overall, 16% of women who reported douching, compared with 10% of those who did not, also reported a history of PID.[12]

Sparse data exist on cultural determinants of the markedly higher douching level in Blacks. Possible explanations include (1) cultural tradition handed down between generations,[2,13] (2) a greater belief in douching as an effective contraceptive practice, and (3) relatively greater marketing of commercial douches in Black communities. Nor do our findings on douching practices help us choose among alternative cultural explanations for higher douching levels among Blacks. We found that Black women douche more frequently; more of them douche after sexual intercourse, and they douche sooner after intercourse, than do Whites.

Our multivariate analysis revealed that neither income nor number of lifetime sex partners were predictors of vaginal douching. We believe this is an artifact of multicollinearity in the data. Income and education are so highly correlated that when the effect of one is statistically eliminated, it is difficult to show the effect of the other.[14] Education was entered into the regression model at the same time as income, and had a masking influence on the latter. Similarly, age at first intercourse and number of lifetime sex partners are highly correlated.

Our findings agree in many ways with the results of earlier studies. For example, a high prevalence of vaginal douching,[1] especially among Blacks,[1,9] has been reported by others. Conversely, some of our findings disagree with earlier reports. In our study, women residing in the western United States had the lowest prevalence of douching; in an earlier report, women from that region had the highest douching prevalence.[1,10] However, previously reported patterns of douching practices were based on surveys of nonrepresentative samples, and are therefore subject to possible measurement errors. Since our data are based on a nationally representative random sample of women of reproductive age, it is unlikely that our findings are affected by sampling bias.

Our finding that vaginal douching is positively associated with a history of PID but negatively associated with a history of gonorrhea, chlamydia trachomatis, genital herpes, or genital warts is consistent with earlier findings.[1,3] If true, this finding suggests that the association between PID and douching is not an artifact of STD, whereby women with STD douche to clear vaginal symptoms, as has been suggested.[15] Rather, despite the lower prevalence of vaginal douching among those who have ever had an STD, women with a history of PID are more likely to douche.

The correlation between douching and PID is also consistent with studies that show strong independent associations between PID sequelae and douching practices.[6,7] Ectopic pregnancy remained correlated with douching even after history of STD and PID[6] and serologic evidence of previous chlamydial infection[3] were controlled. Evidence on the relationship between PID-related tubal infertility and douching is sparse. Douching after menses appears unrelated to risk for primary infertility associated with endometriosis.[16] However, douching has been correlated with an increased risk for cervical carcinoma.[17]

Despite many consistencies among different investigations, the role of douching in the etiology of PID remains difficult to unravel. Plausible associations exist between douching and exposure to sexually transmitted pathogens, acquisition of infection when exposed, development of symptoms, persistence of symptoms and related health care seeking, diagnosis, and development of PID. Moreover, many of these associations are in opposite directions. For example, earlier studies suggested that douching might be useful in preventing and treating vaginal infections.[18–20] Thus, the association between douching and PID may result from (or may be masked by) confounding by sexual, health care seeking, or health provider behaviors.

Because douching is such a common practice in American women of reproductive age and because it may represent a modifiable risk factor for PID and its sequelae, we feel it should be a high priority for future etiologic research. Historical cohort studies, if possible, would provide the most rapidly obtained confirmatory data. These studies might also explore whether any particular aspect of douching practice, such as timing, frequency, or solution used, is most harmful. In addition, reasons for the racial disparity in douching practices need to be investigated. If douching emerges as a true risk factor for PID, this knowledge will be essential in guiding the content of health education campaigns. □

References

1. Forrest KA, Washington AE, Daling JE, Sweet RL. Vaginal douching as a possible risk factor for pelvic inflammatory disease. *J Natl Med Assoc.* 1989;81:159–165.
2. Chacko MR, McGill L, Johnson T, Smith PB, Nenney SW. Vaginal douching in teenagers attending a family planning clinic. *J Adolesc Health Care.* 1989;10:217–219.
3. Wolner-Hanssen P, Eschenbach DA, Paavonen J, et al. Association between vaginal douching and acute pelvic inflammatory disease. *JAMA.* 1990;263:1926–1941.
4. Neumann HH, DeCherney A. Douching and pelvic inflammatory disease. *N Engl J Med.* 1976;295:789. Letter.
5. McGregor JA, Spencer NE, French JI, et al. Psychosocial and behavioral risk factors for acute salpingitis. Presented at the Sixth Annual Meeting of the ISSTDR; August 1, 1985; Brighton, England.
6. Chow WH, Daling JR, Weiss NS, Moore DE, Sodestrom R. Vaginal douching as a

potential risk factor for tubal ectopic pregnancy. *Am J Obstet Gynecol.* 1985;153: 727–729.
7. Chow JM, Yonekura L, Richwald GA, Greenland S, Sweet RL, Schachter J. The association between *Chlamydia trachomatis* and ectopic pregnancy: a matched pair, case-control study. *JAMA.* 1990;263:3164–3167.
8. Cates W, Jr, Rolfs RT, Jr, Aral SO. Sexually transmitted diseases, pelvic inflammatory disease and infertility: an epidemiologic update. *Epidem Rev.* In press.
9. Stock RJ. A review of trends in vaginal douching: the relationship to changing incidences of PID and tubal ectopic pregnancy. *Am J Gynecol Health.* 1989;3:155–161.
10. Stock RJ, Stock ME, Hutto JM. Vaginal douching: current concepts and practices. *Obstet Gynecol.* 1973;42:141–146.
11. Bachrach CA, Horn MC, Mosher WD, Shimizu I. *National Survey of Family Growth, Cycle III: Sample Design, Weighting and Variance Estimation.* National Center for Health Statistics; September 1985. Vital and Health Statistics, Series 2, No. 98.
12. Aral SO, Mosher WD, Cates W, Jr. Self reported pelvic inflammatory disease in the U.S.: 1988. *JAMA.* In press.
13. Snow LF. Traditional health beliefs and practices among lower class black Americans. *West J Med.* 1983;139:820–828.
14. Aral SO, Soskolne V, Joesoef RM, O'Reilly KR. Sex partner selection as risk factor for STD: clustering of risky modes. *Sex Transm Dis.* In press.
15. Dan BB. Sex, lives, and chlamydia rates. *JAMA.* 1990;263:3191–3192.
16. Cramer DW, Wilson E, Stillman RJ, et al. The relation of endometriosis to menstrual characteristics, smoking and exercise. *JAMA.* 1986;255:1904–1908.
17. Gardner J, Schuman K, Slattery N, Sanborn J, Abbot T, Overall J. Is vaginal douching related to cervical carcinoma? Presented at the 23rd annual meeting of the Society for Epidemiologic Research; June 12–15, 1990.
18. Edwards WM. A study of progonosyl using prostitutes in Nevada's legal houses of prostitution. *J Reprod Med.* 1970;11:81.
19. Limsuwan A, Vachrotai S, Panupornpraponng Y, et al. A clinical trial of a vaginal preparation regimen for the prophylaxis of gonorrhea. *J Med Assoc Thai.*
20. Beaton JH, Gibson F, Roland M. Short-term use of a medicated douche preparation in the symptomatic treatment of minor vaginal irritation, in some cases associated with infertility. *Int J Fertil.* 1984;29:109–112.

15th Annual National Conference on Rural Health to Be Held in DC

The National Rural Health Association will hold its 15th annual national conference on rural health, "Rural Health: Caring for the Country," May 6 to 9, 1992, at the Hyatt Regency Crystal City Hotel in Washington, DC.

The conference is designed to

- Develop practical skills and techniques in health services administration, organization, and research
- Present practical clinical sessions of interest to rural health providers
- Showcase current policy issues affecting rural health services
- Report the results of research applicable to rural health service administrators and providers

The conference program includes exhibits as well as award presentations recognizing outstanding achievement of rural health care providers, practices, programs, and research. To receive registration information, contact Robert Quick, 301 East Armour Boulevard, Suite 420, Kansas City, MO 64111. (Tel: 816/756-3140.)

Co-sponsors of the conference include the Center on Rural Elderly at the University of Missouri–Kansas City, the National Clearinghouse for Primary Care Information, the National Rural and Small Schools Consortium, and the National Rural Health Network (a subsidiary of the National Rural Electric Cooperative Association). Conference cohosts are the American Society of Hospital Pharmacists, the American College of Preventive Medicine, and the National Association of County Health Officials.

Employment Status and Heart Disease Risk Factors in Middle-Aged Women: The Rancho Bernardo Study

Donna Kritz-Silverstein, PhD, Deborah L. Wingard, PhD, and Elizabeth Barrett-Connor, MD

ABSTRACT

Background. In recent years, an increasing number of women have been entering the labor force. It is known that in men, employment is related to heart disease risk, but there are few studies examining this association among women.

Methods. The relation between employment status and heart disease risk factors including lipid and lipoprotein levels, systolic and diastolic blood pressure, fasting and postchallenge plasma glucose and insulin levels, was examined in 242 women aged 40 to 59 years, who were participants in the Rancho Bernardo Heart and Chronic Disease Survey. At the time of a follow-up clinic visit between 1984 and 1987, 46.7% were employed, primarily in managerial positions.

Results. Employed women smoked fewer cigarettes, drank less alcohol, and exercised more than unemployed women, but these differences were not statistically significant. After adjustment for covariates, employed women had significantly lower total cholesterol and fasting plasma glucose levels than unemployed women. Differences on other biological variables, although not statistically significant, also favored the employed women.

Conclusions. Results of this study suggest that middle-aged women employed in managerial positions are healthier than unemployed women. (*Am J Public Health.* 1992;82:215–219)

Introduction

Heart disease is the leading cause of death among both men and women.[1,2] In men, risk of heart disease is related to employment.[3,4] However, although more and more women have been entering the labor force in recent years,[5–7] few studies have examined employment and heart disease risk factors in women, and their findings have been inconsistent.[3,4,8–10].

Among young women in the Framingham Offspring Study, homemakers and those employed outside the home were compared and showed no significant differences in blood pressure, total cholesterol, high-density lipoprotein (HDL) level, or low-density lipoprotein (LDL) level over an 8-year period.[10] In contrast, Hazuda et al.[9] found highly significant differences in HDL and in ratio of HDL to total cholesterol, triglycerides, and fasting and postchallenge plasma glucose favoring employed women over homemakers. These differences were not explained by age, socioeconomic status, obesity, exercise, smoking, alcohol consumption, or use of estrogen.

We studied the association of employment status and heart disease risk factors in middle-aged White women. We examined differences between employed and nonemployed women in blood pressure, lipid and lipoprotein levels, and fasting and postchallenge plasma glucose and insulin levels before and after adjusting for potentially confounding physical and lifestyle covariates.

Methods

The population consisted of women who were initially enrolled in the Rancho Bernardo Heart and Chronic Disease Survey between 1972 and 1974 and who attended a follow-up clinic visit between 1984 and 1987; 84% of surviving women aged 40 and older responded. This report is based on the 242 women aged 40 through 59 at the follow-up visit who had fasted the requested 12 hours and were not currently taking lipid-lowering medications. This age group was chosen for analysis because there were too few women over age 59 who were still working. Only one woman was excluded for not fasting the requested 12 hours, and three women were excluded because they were using lipid-lowering medications.

Data Collection

A standardized interview was used to obtain information on current employment status, marital status, cigarette smoking, alcohol consumption (number of drinks per week of beer, wine, or liquor converted to milliliters of alcohol per week), and exercise (whether subjects exercised three or more times per week). Reported current use of estrogen and other medications was confirmed by prescriptions or pill containers brought to the clinic for that purpose.

In the clinic, height (meters), weight (kilograms), and waist and hip circumferences (centimeters) were measured with

The authors are with the Department of Community and Family Medicine, Division of Epidemiology, at the University of California, San Diego.

Requests for reprints should be sent to Elizabeth Barrett-Connor, MD, Division of Epidemiology, Department of Community and Family Medicine, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0607.

This paper was submitted to the journal January 31, 1991, and accepted with revisions September 16, 1991.

TABLE 1—Age-Adjusted Comparisons of Employed and Unemployed Women Aged 40–59 on Socioeconomic Variables: Rancho Bernardo, Calif, 1984–1987 (n = 242)

Variable	Unemployed (n = 113)	Employed (n = 129)	χ^2
	Percent	Percent	
Married (% yes)	92.5	79.0	6.86**
Hollingshead Index[a]			
I (professionals, executives)	48.2	37.0	
II (managers, business owners)	23.2	38.6	
III (administrators)	17.0	16.5	
IV–X (secretarial, sales, clerical, other workers)	11.6	7.9	0.19
Education			
Graduate or professional	1.8	8.7	
College graduate	13.4	22.0	
Some college	42.9	33.9	
High school graduate	38.4	33.9	
Some high school or less	3.5	1.5	11.54*

[a]Hollingshead Index to assess social class of household was based on the usual occupation of the head of household.
*$P < .05$, **$P < .01$.

the participants in light clothing and without shoes. Body mass index (weight/height^2) and waist-hip ratio (waist circumference/hip circumference × 100) were used to estimate obesity and central adiposity, respectively.

An oral glucose tolerance test with a 75-g glucose load was performed between 7 AM and 11 AM after a 12-hour fast. Blood was obtained by venipuncture before and 2 hours after the glucose challenge.

Lipids and lipoproteins were measured in a Lipid Research Clinics laboratory under the continuous standardization program of the Centers for Disease Control in Atlanta, Ga. Fasting plasma total cholesterol and triglyceride levels were measured by enzymatic techniques with an ABA–200 biochromatic analyzer (Abbott Laboratories, Irving, Tex), a high-performance cholesterol reagent (No. 236691, Boehringer-Mannheim Diagnostics, Indianapolis, Ind), and a triglyceride agent (No. 6097, Abbot Laboratories). HDL cholesterol was measured by precipitating the other lipoproteins with heparin and manganase chloride according to the standardized procedures of the Lipid Research Clinics manual.[11] LDL cholesterol was calculated using the Friedwald formula.[12]

Plasma glucose was measured with a glucose oxidase method, and plasma insulin was measured in microunits per milliliter by radioimmunoassay in a diabetes research laboratory.[13] Insulin levels were available for subjects seen after November 4, 1985 (n = 216 for fasting plasma insulin, n = 205 for postchallenge plasma insulin).

Two blood pressure readings taken at least 30 seconds apart were obtained with a standard mercury sphygmomanometer after the participants had been seated for 5 minutes. Categorical hypertension was defined as the average of two systolic blood pressure readings equal to or greater than 160 mm Hg; or the average of two diastolic blood pressure readings equal to or greater than 95 mm Hg; or the current use of antihypertensive medication.

Data Analysis

The distributions of covariates (age, obesity, alcohol consumption, estrogen use, exercise, smoking, marital status, education, and social class, heart disease risk factors, total cholesterol, HDL, LDL, triglycerides, systolic and diastolic blood pressures, and fasting and postchallenge plasma glucose and insulin levels) were calculated separately for women working and not working outside the home.

Comparisons were made with t tests for continuous variables and with chi-square analyses for categorical variables. Comparisons were adjusted for age by analysis of covariance for continuous variables and by the Mantel-Haenzel Extension Test for categorical variables. Multiple regression analyses were used to examine the independent relation of current work status to lipid and lipoprotein levels, blood pressure, and plasma glucose and insulin levels after adjusting for other potentially confounding physical, sociodemographic, and life-style covariates. All statistical tests were two-tailed.

Results

Among this sample of women, 129 (53.3%) were employed and 113 (46.7%) were not employed outside the home at the time of the interview. Of the former group, 89 were primarily employed as managers, administrators, or business owners, Hollingshead Index[a] II and III; 9 were classified as professionals or executives, Hollingshead Index I.[14]

On average, employed women were 1.7 years younger than unemployed women ($t = 3.11$, $P < .01$). As can be seen in Table 1, after adjustment for age, employed women had a significantly higher educational level than unemployed women ($P = .04$) and were significantly less likely to be currently married ($P < .01$). Based on the Hollingshead Index of the usual occupation of the head of the household, there was no significant difference between employed and unemployed women in social class.

Although the differences were not statistically significant, employed women had a lower mean body mass index and waist-hip ratio, consumed less alcohol, were less likely to smoke cigarettes, and were more likely to exercise than unemployed women (see Table 2).

Comparisons of employed and unemployed women on other heart disease risk factors are shown in Table 3. Almost all differences favored employed women. After adjustment for age, currently employed women had significantly lower mean total and LDL cholesterol, triglyceride, and postchallenge insulin levels and lower systolic blood pressure than unemployed women. Currently employed women also had a significantly lower rate of categorically defined or treated hypertension than unemployed women. Additionally, employed women had lower fasting plasma glucose levels than unemployed women, a difference of borderline statistical significance ($P < .07$). HDL levels did not differ by employment status.

Hierarchical multiple regression analyses were performed to examine differences in heart disease risk factors between employed and unemployed women after adjusting for the potentially confounding attributes as shown in Table 4. After adjustment, employed women had total cholesterol levels that were 11.8

[a]Hollingshead Index: I = professionals, executives; II = managers, small-business owners; III = administrators; IV–X = secretaries, sales, and clerical workers.

mg/dL lower ($P < .05$) and fasting plasma glucose levels that were 4.5 mg/dL lower ($P < .05$) than unemployed women. Currently employed women also had a lower rate of hypertension, systolic blood pressures that were 3.8 mm Hg lower, and postchallenge plasma glucose levels that were 9.5 mg/dL lower than unemployed women, differences that were of clinical but not conventional statistical significance (Ps < .10). All other risk factors also favored the employed women although differences were not statistically significant. Analyses using waist-hip ratio instead of body mass index yielded similar results (data not shown).

The possibility that these results were due to a form of self-selection bias known as the healthy worker effect, whereby only the healthiest sought, secured, and maintained jobs, was tested in several ways. Comparisons were made between currently employed and unemployed women on diseases and risk factors ascertained at their baseline enrollment visit between 1972 and 1974. At enrollment, none of these women reported having had a heart attack, heart failure, or a stroke. They would have been 30 to 49 years old at that time—an age before heart or other chronic disease commonly affects the ability to work. *All* four women who reported diabetes at baseline were employed at the follow-up (1984 to 1987) visit. Regression analyses restricted to women with "normal" risk factor levels at baseline (i.e., total cholesterol below 225 mg/dL; systolic blood pressure below 140 mm Hg; fasting plasma glucose below 140 mg/dL) indicated that currently employed women had lower risk factors at follow-up (total cholesterol levels 12.9 mg/dL lower [$F = 3.75$, $P < .05$], systolic blood pressure 4.1 mm Hg lower [$F = 2.90$, $P < .08$], and fasting plasma glucose levels 5.2 mg/dL lower [$F = 5.88$, $P < .01$]) than unemployed women. Overall, these data do not support selection bias as an explanation for the more favorable risk factor status among the currently employed women.

Discussion

In the present study, women working outside the home were generally healthier than unemployed women. They had significantly lower total cholesterol and fasting plasma glucose levels than women not currently employed. Differences on all other cardiovascular disease risk factors, although not statistically significant, also favored employed women. Furthermore,

TABLE 2—Age-Adjusted Comparisons[a] of Employed and Unemployed Women Aged 40–59 on Physical and Life-style Variables: Rancho Bernardo, Calif, 1984–1987 (n = 242)

Variable	Unemployed (n = 113)	Employed (n = 129)	*F* or χ^2
	Mean	Mean	
Body mass index (kg/m²)	24.9	24.4	0.72
Waist-hip ratio	77.7	76.9	0.84
Alcohol (mL/wk)	90.9	75.6	1.02
	Percent	Percent	
Using estrogen (% yes)	35.1	35.3	0.01
Regular exercise (% yes)	22.2	29.8	1.48
Smoking (% yes)	39.3	33.5	0.67

[a]Comparisons were adjusted for age by analysis of covariance for the continuous variables of body mass index, waist-hip ratio, alcohol consumption, and by the Mantel-Haenzel Extension Test for the categorical variables of estrogen use, exercise and cigarette smoking.

TABLE 3—Age-Adjusted Comparisons[a] of Employed and Unemployed Women Aged 40–59 on Heart Disease Risk Factors: Rancho Bernardo, Calif, 1984–1987 (n = 242)

Risk Factor	Unemployed (n = 113)	Employed (n = 129)	*F* or χ^2
Cholesterol (mg/dL)	226.8	216.6	4.63**
High-density lipoprotein (mg/dL)	68.4	68.0	0.03
Low-density lipoprotein (mg/dL)	135.5	128.5	2.24
Triglycerides (mg/dL)[a]	97.7	85.1	3.76**
Fasting plasma glucose (mg/dL)	99.1	95.4	3.46*
Postchallenge plasma glucose (mg/dL)	120.6	114.3	1.43
Fasting insulin (μ/dL) (n = 216)[a]	12.3	11.2	1.61
Postchallenge insulin (μ/dL) (n = 205)[a]	66.1	53.7	3.82**
Systolic blood pressure (mm Hg)	123.9	118.8	5.24**
Diastolic blood pressure (mm Hg)	76.2	74.9	1.16
Hypertension (% yes)	27.6	12.8	7.02***

[a]Statistics computed on $\log_{10}$ triglycerides, $\log_{10}$ fasting insulin, and $\log_{10}$ nonfasting insulin, respectively.
[b]Age-adjusted comparisons of hypertension rates were computed with the Mantel-Haenzel Extension Test. Age-adjusted comparisons on all other risk factors were computed with analysis of covariance.
*$P < .07$, **$P < .05$, ***$P < .01$.

TABLE 4—Mean Differences between Employed and Unemployed Women in Heart Disease Risk Factors after Adjustment for Possible Confounders:[a] Rancho Bernardo, Calif, 1984–1987

Risk Factor	Difference[b]	*F*
Cholesterol (mg/dL)	−11.8	4.57**
High-density lipoprotein (mg/dL)	0.9	0.10
Low-density lipoprotein (mg/dL)	−7.9	2.19
Triglycerides (mg/dL)	−1.1	1.88
Fasting plasma glucose (mg/dL)	−4.5	5.78**
Postchallenge plasma glucose (mg/dL)	−9.5	3.05*
Fasting plasma insulin (μ/dL)	−1.0	0.22
Postchallenge plasma insulin (μ/dL)	−1.2	2.03
Systolic blood pressure (mm Hg)	−3.8	2.80*
Diastolic blood pressure (mm Hg)	−0.7	0.28
Hypertension (% yes)	−10.1	3.49*

[a]Possible confounders are age, body mass index, estrogen use, alcohol consumption, cigarette smoking, regular exercise, marital status, education, and social class.
[b]Difference is the beta weight obtained from the multiple regression analysis in which unemployed = 1 and employed = 2.
*$P < .10$, **$P < .05$.

these differences were not explained by age, obesity, estrogen use, alcohol consumption, cigarette smoking, exercise, marital status, education, or socioeconomic status.

The results of the present study are in accord with those of Hazuda et al.,[9] who found more favorable lipid levels among employed Mexican-American and non-Hispanic White women after adjusting for age, obesity, socioeconomic status, smoking, exercise, alcohol use, and use of oral contraceptives. Hazuda et al.[9] also found that fasting and postchallenge plasma glucose levels differed between employed and unemployed Mexican-American women but not between employed and unemployed non-Hispanic White women, a difference that was explained by greater obesity in the former ethnic group.[9] In contrast, the present study of upper-middle-class White women, who are on average leaner than other US women,[15] found clinically significant differences in plasma glucose levels between employed and unemployed women, independent of obesity.

The results of the present study are in accord with several studies that demonstrated better self-reported health among employed women.[16-19] However, they are in contrast to results of the Framingham Heart Study,[3,8] which found employment status to have no effect on women's self-reported health. This difference may reflect the Framingham definition of working, according to which any woman who had worked outside her home for more than half her adult life was classified as employed, regardless of her current employment status. The Rancho Bernardo study considered a woman employed only if she was currently employed, the same definition used in other studies.[4,9]

Two studies have reported that the beneficial effects of working on self-reported health are greater for unmarried women than for married women.[20,21] In the present study, where over three quarters of the women were married and there were only nine single unemployed women, adjusting for marital status did not materially alter the results.

Differences did not appear to be due to the self-selection bias called the healthy worker effect, whereby only the healthiest are employed outside the home.[22] The differences between employed and unemployed women remained even after the analyses were restricted to those who had the most optimal risk factors at enrollment. Women with diabetes, the only chronic disease reported at baseline, were all employed at the time of their follow-up visit.

Differences between employed and unemployed women were not explained by differences in socioeconomic status or educational level. Furthermore, given the homogeneous, upper-middle-class status of this cohort, it is unlikely that the differences reflect differential access to medical care.

Employment could be associated with health through other life-style and behavioral factors. Although not statistically significant, differences between employed and unemployed women in obesity, alcohol consumption, cigarette smoking, estrogen use, and regular exercise tended to favor a more healthy life-style among employed women, and adjustment for these attributes reduced the differences in heart disease risk factors between these women (Table 4).

A psychosocial factor could also moderate the relationship between employment status and health. Most of these upper-middle-class women were married and were probably working for job satisfaction rather than financial need. Several studies have found that the association between employment and health is stronger for women who have a more favorable attitude toward their jobs (See Waldron and Herold 1986[23] for review).

A positive job attitude could parallel the perception of control offered by the occupational setting.[24] Individuals who feel that what happens to them is the result of factors beyond their control are said to have an external locus of control, whereas those who believe that what happens to them is the result of their own decisions are said to have an internal locus of control.[25] Those with an internal locus of control are reported to experience less illness,[26] less coronary disease,[27] less hypertension,[28] and less depression[29] than those with an external locus of control. In terms of life-style factors, those with an internal locus of control are less likely to smoke cigarettes[30] and more likely to engage in physical exercise.[31] Middle-aged employed women have been found to have a greater sense of personal control over their environment than women who are homemakers or students,[24] possibly reflecting their income and job responsibilities.

Whatever the mechanism, the present study suggests that upper-middle-class women working outside the home have more favorable heart disease risk factors than unemployed women. Given that almost 80% of the employed women studied were working as professionals, managers, or administrators, these data do not support prior predictions that, as women attain male professional- or executive-level occupations, they will also acquire male coronary heart disease risk and mortality rates.[32] □

Acknowledgments

This work was supported by grants AG07181 from the National Institute of Aging and DK301801 from the National Institute of Diabetes and Digestive and Kidney Diseases.

References

1. Thom TJ, Epstein FH, Feldman JJ, Leaverton PE. Trends in total mortality and mortality from heart disease in 26 countries from 1950 to 1978. *Int J Epidemiol.* 1985;14:510–520.
2. Thom TJ. Cardiovascular disease mortality among United States women. In: Eaker E, Packard B, Wenger NK, Clarkson TB, Tyroler H, eds. *Coronary Heart Disease in Women.* New York, NY: Haymarket-Doyma; 1988:33–41.
3. Haynes SG, Feinleib M. Women, work and coronary heart disease: prospective findings from the Framingham Heart Study. *Am J Public Health.* 1980;70:133–141.
4. House JS, Strecher V, Metzner HL, Robbins CA. Occupational stress and health among men and women in the Tecumseh Community Health Study. *J Health Soc Behav.* 1986;27:62–77.
5. National Commission on Working Women. *An Overview of Women in the Work Force.* Washington, DC: National Commission on Working Women; 1987.
6. US Department of Labor, Bureau of Labor Statistics. *Employment and Earnings.* Vol 35(2). Washington, DC: US Govt Printing Office; December 1978.
7. Mathews KA, Rodin J. Women's changing work roles: impact on health, family, and public policy. *Am Psychol.* 1989;44:1389–1393.
8. Haynes SG, Eaker ED, Feinleib M. The effect of employment, family, and job stress on coronary heart disease patterns in women. In: Gold EB, ed. *The Changing Risk of Disease in Women: An Epidemiological Approach.* Lexington, Mass: DC Health; 1984:37–48.
9. Hazuda HP, Haffner SM, Stern MP, Knapp JA, Eifler CW, Rosenthal M. Employment status and women's protection against coronary heart disease. *Am J Epidemiol.* 1986;123:623–640.
10. Hubert HB, Eaker D, Garrison RJ, Castelli WP. Life-style correlates of risk factor change in young adults: an eight-year study of coronary heart disease risk factors in the Framingham offspring. *Am J Epidemiol.* 1987;125:812–831.
11. Lipid Research Clinics Program. *Manual of Laboratory Operations.* Vol 1. *Lipid and Lipoprotein Analysis.* 2nd ed. Washington, DC: US Govt Printing Office; 1974. US Dept of Health, Education, and Welfare publication NIH 75–628.
12. Friedwald WJ, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma without use of the preparative ultracentrifuge. *Clin Chem.* 1972;18:459–502.
13. Desbuquois B, Ausbach GD. Use of poly-

ethylene glycol to separate free and antibody-bound peptide hormones in radioimmunoassay. *J Clin Endocrinol Metab.* 1971;33:732–738.
14. Hollingshead AB. *Two-Factor Index of Social Position.* Privately published; 1957.
15. Barrett-Connor E. The prevalence of diabetes mellitus in an adult community as determined by history of fasting hyperglycemia. *Am J Epidemiol.* 1980;111:705–712.
16. Adelman PK, Antonucci TC, Crohan SE, Coleman LM. A causal analysis of employment and health in midlife women. *Women Health.* 1990;16:5–20.
17. Jennings S, Mazaik C, McKinlay S. Women and work: an investigation of the association between health and employment status in middle-aged women. *Soc Sci Med.* 1984;19:423–431.
18. Verbrugge L. Role burdens and physical health of women and men. *Women Health.* 1986;11:47–77.
19. Waldron I. Employment and women's health: an analysis of causal relationships. *Int J Health Serv.* 1980;10:435–454.
20. Waldron I, Jacobs J. Effects of labor force participation on women's health: new evidence from a longitudinal study. *J Occup Med.* 1988;30:977–983.
21. Repetti RL, Mathews KA, Waldron I. Employment and women's health: effects of paid employment on women's mental and physical health. *Am Psychol.* 1989;11: 1394–1401.
22. McMichael AJ, Haynes SG, Tyroler HA. Observations of the evaluation of occupational mortality data. *J Occup Med.* 1975;17:128–131.
23. Waldron I, Herold J. Employment, attitudes toward employment and women's health. *Women Health.* 1986;11:79–98.
24. Erdwins CJ, Mellinger JC. Midlife women: relation of age and role to personality. *J Pers Soc Psychol.* 1984;47:390–395.
25. Rotter JB. Generalized expectancies for internal versus external control of reinforcement. *Psychol Monogr.* 1966;80(1, whole no. 609).
26. Kobasa SC. Stressful live events, personality and health: an inquiry into hardiness. *J Pers Soc Psychol.* 1979;37:1–11.
27. Cromwell RL, Butterfield EC, Brayfield FM, Curry JL. *Acute Myocardial Infarction: Reaction and Recovery.* St. Louis, Mo: CV Mosby; 1977.
28. Naditch MP. Locus of control, relative discontent and hypertension. *Soc Psychiatry.* 1974;9:111–117.
29. Lefcourt HM. *Locus of Control: Current Trends in Theory and Research.* Hillsdale, NJ: Lawrence Erlbaum; 1976.
30. Coan RW. Personality variables associated with cigarette smoking. *J Pers Soc Psychol.* 1973;26:86–104.
31. Walker TG. Behavior of temporary members in small groups. *J Appl Psychol.* 1973;58:144–146.
32. Baruch GK, Biener L, Barnett RC. Women and gender in research on work and family stress. *Am Psychol.* 1987;42: 130–136.

AIDS in Bisexual Men in the United States: Epidemiology and Transmission to Women

ABSTRACT

Background. Homosexual and bisexual men with acquired immunodeficiency syndrome (AIDS) differ, and bisexual men play an important role in the sexual transmission of human immunodeficiency virus (HIV) to women.

Methods. To describe AIDS in these groups, we examined AIDS cases reported nationally through June 1990.

Results. Among 65 389 men who reported having had sex with men since 1977, 26% were bisexual. More Black (41%) and Hispanic men (31%) than White men (21%) reported bisexual behavior. Bisexual men were twice as likely to report intravenous drug use (20%) as were homosexual men (9%), regardless of race or ethnicity. Among 3555 women with heterosexually acquired AIDS, 11% reported sexual contact with a bisexual man and no other risk factor, although in some states approximately half reported such contact. In 1989, the AIDS rate due to sex with a bisexual man was three and five times higher among Hispanic and Black women, respectively, than among White women.

Conclusions. Differences between bisexual and homosexual men with AIDS and the relative importance of AIDS in women due to sexual contact with bisexual men should be considered in the development of HIV prevention programs. (*Am J Public Health.* 1992;82:220–224)

Susan Y. Chu, PhD, Thomas A. Peterman, MD, Lynda S. Doll, PhD, James W. Buehler, MD, and James W. Curran, MD

Introduction

Most reports on human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS) consider bisexual and homosexual men as a single group. However, this categorization may not always be appropriate. Bisexual men may differ from homosexual men in several behavioral and epidemiologic aspects of HIV infection[1,2] and these differences have important public health implications. Interventions designed to prevent HIV transmission among homosexual men may not reach or be appropriate among bisexual men. For example, bisexual men may be less likely to identify with gay communities or to view themselves as at risk for HIV infection.[3,4] In addition, bisexual men may transmit HIV to their female partners.

In this study, we analyzed national AIDS surveillance data to compare bisexual men (men who reported sexual contact with both men and women) and homosexual men (men who reported sexual contact with men and not with women) with AIDS and to assess the relative importance of bisexual men in the heterosexual transmission of HIV to women.

Methods

We used national surveillance data for cases of AIDS in persons over 13 years of age who were reported to the Centers for Disease Control (CDC) from June 1981 through June 1990.

The sexual behavior of men was determined by using responses to the following two questions on the AIDS case report form:

> "After 1977 and preceding the diagnosis of AIDS, did this patient have sexual relations with a male partner?"
> "After 1977 and preceding the diagnosis of AIDS, did this patient have sexual relations with a female partner?"

Men were classified as homosexual if they were reported as having had sex with men and not with women, and were classified as bisexual if they were reported as having had sex with both men and women. Men who were reported as having had sex with men but for whom information on sex with women was missing were excluded.

Data on AIDS indicator diseases were also obtained from AIDS case reports; more than one condition can be specified. The surveillance definition for AIDS was revised in September 1987 to include additional indicator diseases and to accept the presumptive diagnoses of some other indicator diseases.[5] Persons meeting only the new criteria were more likely to be Black, Hispanic, or intravenous (IV) drug users[6]; therefore, we repeated our analyses, including only patients who met the pre-1987 definition. Because results were similar, we report findings based on all reported cases meeting the current case definition. Analyses of indicator conditions added with the revised case definition (e.g., wasting syndrome) were restricted to persons reported after September 1987.

A woman was classified as having AIDS due to heterosexual transmission if

Susan Y. Chu, Lynda S. Doll, James W. Buehler, and James W. Curran are with the Division of HIV/AIDS, National Center for Infectious Diseases, and Thomas A. Peterman is with the Division of Sexually Transmitted Diseases, National Center for Prevention Services, all at the Centers for Disease Control, Atlanta, Ga.

Requests for reprints should be sent to Susan Y. Chu, PhD, Surveillance Branch, Division of HIV/AIDS, Center for Infectious Diseases, Centers for Disease Control, Mailstop E-47, 1600 Clifton Rd NE, Atlanta, GA 30333.

This paper was submitted to the journal May 22, 1991, and accepted with revisions August 22, 1991.

Editor's Note. See related editorial by Lifson on page 166 of this issue.

she had no history of blood transfusion or IV drug use and reported having had sexual contact with a person who had AIDS or HIV infection, or with a person who had any of the following risks since 1977: IV drug use, male-to-male sexual contact, or receipt of blood or a blood product.[7] Because we were primarily interested in HIV transmission in the United States, we excluded women born in countries where heterosexual contact is the primary means of HIV transmission (i.e., African and Caribbean countries). To assess the relative importance of bisexual men in the heterosexual transmission of HIV among women, we compared the number of cases of AIDS in women attributed to sexual contact with a bisexual man with the number of cases attributed to sexual contact with a male IV drug user. We used sex- and race-specific mid-year (July 1) population estimates from public-use computer tapes prepared from Bureau of Census data to calculate annual AIDS rates.[8] Population estimates for persons of Hispanic origin were extrapolated from US Bureau of the Census reports.[9,10]

Results

Epidemiologic Differences Between Bisexual and Homosexual Men with AIDS

Complete responses to the two sexual history questions were available for 65 389 men who reported having had sex with men since 1977; of these, 16 793 (26%) also reported having had sex with women. The proportion of bisexual men has increased slightly over time, from 23% in 1983 to 26% in 1989.

For an additional 26 092 men who reported having had sex with men, information regarding sexual contact with women was missing. The distribution of characteristics for these men was intermediate to those of men classified as homosexual or bisexual, suggesting that some were bisexual and some were homosexual.

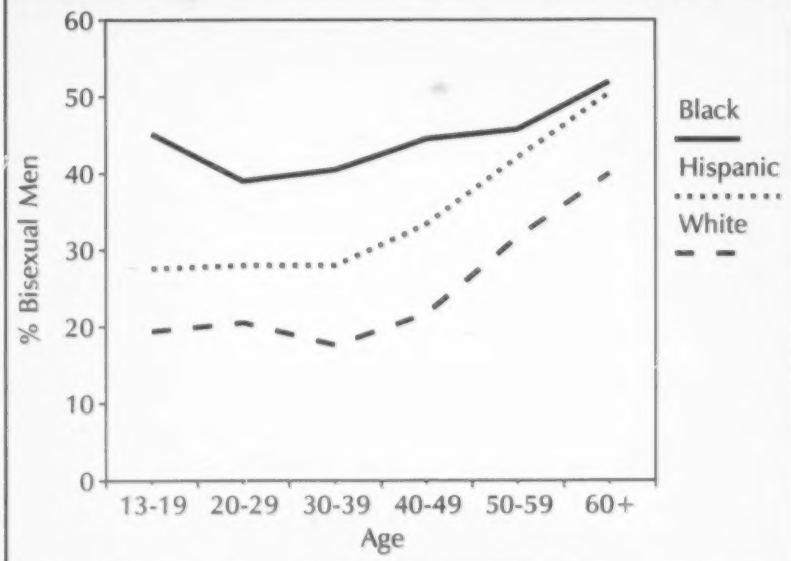
Bisexual and homosexual men differed by race/ethnicity. Among men who reported sex with men (homosexual and bisexual combined), bisexuality was reported by more Black (41%) and Hispanic (31%) men than White men (21%) (Table 1). These differences were fairly consistent over time and in different states.

The proportion of men who were bisexual increased with age after age 30 to 39 years (Figure 1). The pattern with age was consistent among racial/ethnic groups except for young Black men (13–19 years of age). These differences by age were not affected by IV drug use or year of diagnosis.

TABLE 1—Bisexual and Homosexual Men Reported with AIDS, United States, through June 1990

Race/ethnicity	Bisexual	Homosexual	% Bisexual[a]
White	9555	36 260	21
Black	4842	6851	41
Hispanic	2217	5049	31
Asian/Pacific Islander	152	367	29
American Indian/Alaskan Native	27	69	28
Total	16 793	48 596	26

Note. Men were classified as bisexual if they reported having had sex with both men and women since 1977, and were classified as homosexual if they reported having had sex with men and not with women.
[a]% Bisexual = [(number of bisexual men with AIDS)/(number of bisexual men with AIDS + number of homosexual men with AIDS)] × 100.



*Percent bisexual men = $\frac{\text{No. bisexual men with AIDS}}{\text{(No. bisexual + No. homosexual men with AIDS)}} \times 100$

FIGURE 1—Men with AIDS who had sex with men (percent bisexual), by age and race, United States, through June 1990.

Bisexual men with AIDS were twice as likely as homosexual men to report IV drug use (Table 2). The increased use of IV drugs among bisexual men was observed in each racial/ethnic group, with the most pronounced difference among Hispanic men. This difference in IV drug use between bisexual and homosexual men with AIDS was also consistent by region (IV drug use in bisexual men/IV drug use in homosexual men: Northeast, 2.4; Midwest, 2.4; South, 2.2; West, 1.9) and by year of diagnosis.

The opportunistic diseases reported for bisexual and homosexual men with AIDS were similar except for Kaposi's sarcoma and wasting syndrome. Kaposi's sarcoma (including presumptive and definitive diagnoses) was about 1.5 times more common among homosexual than among bisexual men with AIDS (Table 3). This difference persisted across different racial/ethnic groups, although the difference was smaller among Black men, and also was consistent across age groups. Although the frequency of Kaposi's sarcoma in all men with AIDS has declined over time, the difference between homosexual and bisexual men in frequency of Kaposi's sarcoma persisted by diagnosis year. For example, in 1984, 38% of homosexual men and 27% of bisexual men were reported with Kaposi's sarcoma; in 1989, 17% of

TABLE 2—History of Intravenous Drug Use in Bisexual and Homosexual Men Reported with AIDS, United States, through June 1990

	Percentage with History of Intravenous Drug Use[a]			
	White (n = 44 034)	Black (n = 10 521)	Hispanic (n = 6614)	Total[b] (n = 61 784)
Bisexual	16	23	27	20
Homosexual	8	14	11	9

[a]Excludes 1781 White, 1172 Black, and 652 Hispanic men with unknown histories of intravenous drug use.
[b]Includes 519 Asian/Pacific Islanders and 96 Native Americans.

TABLE 3—Kaposi's Sarcoma and HIV Wasting Syndrome in Bisexual and Homosexual Men Reported with AIDS, United States, through June 1990

	White	Black	Hispanic	Total
	Percentage with Kaposi's Sarcoma			
	(n = 45 815)	(n = 11 693)	(n = 7266)	(n = 65 389)
Bisexual	17	8	15	14
Homosexual	23	11	22	21
	Percentage with HIV Wasting Syndrome			
	(n = 35 077)	(n = 9764)	(n = 5754)	(n = 51 095)[b]
Bisexual	19	20	20	19
Homosexual	14	16	11	14

[a]Includes 519 Asian/Pacific Islanders and 96 Native Americans.
[b]Includes persons reported after September 1987 only; includes 410 Asian/Pacific Islanders and 90 Native Americans.

* Excludes 110 women with AIDS who reported sexual contact with both a bisexual man and an IV drug user.

FIGURE 2—Women with AIDS attributed to sexual contact with a bisexual man or sexual contact with IV drug user, United States, through June 1990.

homosexual men and 10% of bisexual men were reported with Kaposi's sarcoma. Findings were similar when IV drug users were excluded from the analyses.

Wasting syndrome was more common among bisexual men with AIDS than among homosexual men with AIDS (Table 3). This difference was consistent regardless of race or IV drug use, except for Black IV drug users, among whom no difference in the frequency of wasting syndrome was seen. This difference in wasting also was consistent by diagnosis year and by region of the United States, although the difference was greater in the US Territories (wasting diagnosis in bisexual men, 46%; in homosexual men, 36% and smaller in the West (wasting diagnosis in bisexual men, 16%; in homosexual men, 14%).

Women with AIDS Reporting Contact with a Bisexual Man

Among 3555 women with AIDS who acquired HIV infection through heterosexual contact (excluding women born in African or Caribbean countries), 11% (n = 405) reported sexual contact with a bisexual man and no other risk factor for AIDS. Another 114 women with AIDS reported sexual contact with a bisexual man and another specific heterosexual contact (an IV drug user, n = 110; a transfusion recipient or person with hemophilia, n = 4).

In 1989, the rate for AIDS due to sex with a bisexual man (and no other partner at risk) was three and five times higher among Hispanic (1.8 per 1 000 000) and Black (2.8 per 1 000 000) women, respectively, than among White women (0.6 per 1 000 000), and was seven times lower overall than the rate for AIDS due to sex with an IV drug user. However, the relative importance of AIDS attributed to sexual contact with bisexual men varied widely by state (Figure 2). In states with a relatively low HIV seroprevalence in IV drug users (e.g., California), the number of cases in women associated with sex with a bisexual man was comparable to the number associated with sex with an IV drug user. In states with high HIV seroprevalence in IV drug users (e.g., New Jersey and New York), AIDS due to sexual contact with an IV drug user accounted for most heterosexually acquired AIDS in women, regardless of race.

Discussion

Our estimate that 14% of all men with AIDS are bisexual is a minimum estimate; if the frequency of bisexuality was the same in the group with missing data as in the group with complete information, bisexual men with AIDS would represent 19% of all men with AIDS. The percentage of men who were bisexual has increased slightly over time. This increase

may be influenced by the sexual history questions, which asked only about sexual contact with men or women since 1977. Reports for more recent cases would cover a longer time period and these questions could not distinguish between early and more recent sexual behavior.

Among men with AIDS who reported having had sex with men, Blacks reported bisexuality more frequently than did Whites or Hispanics. Similar findings have been reported elsewhere. In a study of sexually transmitted disease clinic patients in Ohio, among 1107 men who reported having had sex with other men, 47% of the Black men and 27% of the White men classified themselves as bisexual.[11] In a study conducted at 20 large blood centers, 209 HIV-infected prospective male blood donors reported having had sex with men; 48% of Blacks, 36% of Hispanics, and 33% of Whites had had sex with a woman in the previous year.[3] Some have suggested that the adoption of a bisexual life-style among Black men may be related to the difficulty of being homosexual in the Black community.[1] However, these data do not enable us to distinguish between a difference in reporting and a true difference in behavior by race/ethnicity.

The Hispanic population in the United States is diverse in country of origin and degree of acculturation. Unless bisexual Hispanic men are highly acculturated, their behavior patterns may be more similar to behavioral patterns of their country of origin.[1] From studies of bisexual behavior in Mexico, it was estimated that as many as 30% of Mexican men between the ages of 15 and 25 may engage in sex with both men and women.[12] However, men who practiced the active, insertive role in male-to-male sex and who also had sex with women did not view themselves as homosexual or bisexual; thus they were a potentially difficult group to reach.[12,13] Because of the diversity of the Hispanic population, generalizations about AIDS-related behaviors must be made cautiously.[1]

Older men with AIDS who reported having had sex with men were more likely to report bisexual than homosexual behavior. This fact may represent the greater acceptance of homosexuality during the gay liberation movement of the 1960s and 1970s (a cohort effect), but we cannot exclude other reasons, such as an increasing trend with age or differences in reporting of sexual activity with age. For example, because men were asked whether they had had any male or female partners after 1977 and preceding their diagnosis of AIDS, the time period of sexual activity covered is longer for older men than for younger men, which could increase the likelihood of their having had at least one female partner.

Kaposi's sarcoma was more common in homosexual men than in bisexual men with AIDS. If bisexual men with AIDS had fewer sexual contacts with male partners than did homosexual men, this finding supports suggestions that Kaposi's sarcoma is associated with a sexually transmitted agent that is more prevalent among men who have sex with men.[14] There is some evidence that bisexual men have fewer male partners than do homosexual men. A study of 1034 single men living in San Francisco found that during a 6-month period (January to June 1984), 27% (172/641) of the homosexual men and 21% (37/173) of the bisexual men reported having had more than 10 sexual partners.[15] Preliminary data from a multicenter study of HIV-seropositive blood donors found that men who reported only homosexual contact had had a larger number of male partners since 1978 (median number of male partners = 10, n = 148) than men who reported bisexual contact (median number of male partners = 5, n = 58) (CDC, unpublished data). The hypothesis that Kaposi's sarcoma is associated with a sexually transmitted agent is also supported by the finding that Kaposi's sarcoma was more common in women who reported a bisexual male partner than in women who did not report a bisexual male partner.[14]

The higher frequency of HIV wasting syndrome in bisexual men than in homosexual men was consistently observed across racial/ethnic groups, even after the exclusion of IV drug users. This finding is more difficult to explain. Possible reasons include differences between bisexual and homosexual men with AIDS in time of diagnosis, in completeness of medical evaluation, or in conditions associated with severe weight loss.

These findings suggest that bisexual and homosexual men with AIDS are distinct. For example, more bisexual men than homosexual men reported a history of IV drug use. Similar information was found in a study of 200 male prostitutes, in which self-identified bisexual prostitutes were more likely than homosexual prostitutes to use intravenous cocaine or heroin, to be homeless, and to have "hustled" in a number of large cities.[16] Although these male prostitutes had sex for money with men, more of their other sexual encounters were with women. In preliminary data linking AIDS surveillance and death certificate information, 86% of homosexual men versus 52% of bisexual men had never married. At death, 3% of homosexual men versus 24% of bisexual men were currently married (CDC, unpublished data). Bisexual men who are married or who do not identify themselves as homosexual may be more difficult to reach, especially through gay community networks. Because of these differences, bisexual and homosexual men should be considered separately in certain research and prevention settings.

We used categories based on reported sexual behaviors, not on self-labeled sexual identity. Our "bisexual" category does not necessarily represent men who identify themselves as bisexual, but more accurately represents men with AIDS who were reported as having had sexual contact with both male and female partners since 1977. Self-identification does not necessarily predict sexual behavior, and some men who report or consider themselves to be homosexual or heterosexual actually have sex with both men and women.[1] These differences between sexual identity and sexual behavior must be considered in interpreting these findings and in counseling persons about HIV.

Nationwide, 63% of women with heterosexually acquired AIDS were infected through sexual contact with IV drug using men.[17] Rates of AIDS attributed to sexual contact with IV drug users vary widely by state and race/ethnicity, with rates ranging from 0 in some states to as high as 14.5 per 100 000 among Black women in New Jersey. Approximately 70% of all AIDS cases in women related to sexual contact with an IV drug user are reported from New York, New Jersey, Florida, and Puerto Rico. Rates of AIDS in women attributed to sexual contact with bisexual men vary less by state (ranging from 0 to 0.9 per 100 000 among black women in Florida). In some states where HIV prevalence among IV drug users is low, sexual contact with bisexual men accounts for a substantial proportion of women with heterosexually acquired AIDS.

Nearly one quarter of bisexual men with AIDS who died were married at the time of death. Because many women may not be aware of their partners' behavior,[18,19] the number of women with AIDS who have a bisexual partner is probably higher than reported. Thirteen percent of women with heterosexually acquired AIDS did not report the risk of their partner and 7% of all women with AIDS are reported without an established risk fac-

tor[7]; some of these cases may have resulted from sexual contact with bisexual men. In addition, awareness may differ by race/ethnic group; among 52 female partners of HIV-seropositive bisexual men enrolled in the California Partners' Study, 80% (28/35) of White, 20% (1/5) of Black, and 22% (2/9) of Latino women were aware of their partners' bisexuality at entry into the study.[19] Similar findings were reported from Chicago.[20] Rates for AIDS due to sexual contact with bisexual men were highest among Black and Hispanic women, yet these women may be the ones least aware of their risk. In addition, this difference in awareness would result in a greater underestimation of the rate of AIDS due to sex with bisexual men among Black and Hispanic women.

Although bisexual and homosexual men are often considered together, differences between them should be acknowledged in efforts to prevent both male-to-male and male-to-female HIV transmission. Effective AIDS risk-reduction programs that reach bisexual men who put themselves and their male and female partners at risk for HIV will require a better understanding of the patterns of bisexual behavior. Critical issues include the incidence of sexual behaviors with male and female partners, the extent of disclosure of same-sex behaviors to female partners, the social and sexual networks of bisexuals, cultural or ethnic differences in bisexual behavior, and the relationship between sexual identity and sexual behavior.[1] □

Acknowledgments

We thank Ruth L. Berkelman, MD, Sevgi O. Aral, PhD, Willard Cates, Jr, MD, and George P. Schmid, MD, for their thoughtful comments and review of this manuscript.

References

1. Doll LS, Peterson J, Magana JR, Carrier JM. Male bisexuality and AIDS in the United States. In: R. Tielman, M. Carballo, A. Hendriks, eds. *Bisexuality and HIV/AIDS.* Buffalo, NY: Prometheus Books; 1991.
2. Mayo DJ, Doll LS, Machette S. Behaviorally bisexual men and AIDS. Report from the Centers for Disease Control Conference on Bisexuality; October 1990. Atlanta, Ga: Centers for Disease Control.
3. Doll LS, Petersen LR, White CR, Johnson ES, Ward JW, et al. Homosexually and nonhomosexually identified men who have sex with men: a behavioral comparison. *J Sex Res.* In press.
4. Bossio L, Gilvonio J, Carovano K. Bisexuality and risk behavior in Lima, Peru. Abstract presented at the Sixth International Conference on AIDS; June 1990; San Francisco, Calif.
5. Centers for Disease Control. Revision of the CDC surveillance case definition for acquired immunodeficiency syndrome. *MMWR.* 1987;36(suppl 1S).
6. Selik RW, Buehler JW, Karon JM, Chamberland ME, Berkelman RL. Impact of the 1987 revision of the case definition of acquired immune deficiency syndrome in the United States. *J AIDS.* 1990;3:73–82.
7. *HIV/AIDS Surveillance Report.* Atlanta, Ga: Centers for Disease Control; July 1990.
8. Irwin R. *1980–1989 Intercensal population estimates by race, sex, and age* [machine-readable data file]. Alexandria, Va: Demo-Detail; 1990.
9. *Population Estimates by Race and Hispanic Origin for States, Metropolitan Areas, and Selected Counties: 1980 to 1985.* Washington, DC: US Government Printing Office; 1989. US Bureau of the Census, Current Population Reports, Series P-25, No. 1040-RD-1.
10. *Projections of the Hispanic Population: 1983 to 2080.* Washington, DC: US Government Printing Office; 1986. US Bureau of the Census, Current Population Reports, Series P-25, No. 995.
11. Kramer MA, Aral SO, Curran JW. Self-reported behavior patterns of patients attending a sexually transmitted disease clinic. *Am J Public Health.* 1980;70:997–1000.
12. Carrier JM. Mexican male bisexuality. In: Klein F, Wolf TJ, eds. *Bisexualities: Theory and Research.* New York, NY: Haworth Press; 1985.
13. Hernandez G, Uribe P, Hernandez M, et al. Sexual behavior, HIV status, and condom use in homosexual and bisexual men in Mexico City. Abstract presented at the Seventh International Conference on AIDS; June 1991; Florence, Italy.
14. Beral V, Peterman TA, Berkelman RL, Jaffe HW. Kaposi's sarcoma among persons with AIDS: a sexually transmitted infection? *Lancet.* 1990;335:123–128.
15. Winkelstein W Jr, Samuel M, Padian NS, Wiley JA. Selected sexual practices of San Francisco heterosexual men and the risk of infection by the human immunodeficiency virus. *JAMA.* 1987;257:1470–1471. Letter.
16. Boles J, Elifson K, Sweat M. Self-reported bisexuality among male prostitutes. Presented at the CDC Workshop on Bisexuality and AIDS; October 1989; Atlanta, Ga.
17. Ellerbrock TV, Bush TJ, Chamberland ME, Oxtoby MJ. Epidemiology of women with AIDS in the United States, 1981–1990. *JAMA.* 1991;265:2971–2975.
18. Hays D, Samuels A. Heterosexual women's perceptions of their marriages to bisexual or homosexual men. *J Homosex.* 1989;18:81–100.
19. Padian NS. Female partners of bisexual men. Presented at the CDC Workshop on Bisexuality and AIDS; October 1989; Atlanta, Ga.
20. McKirnan D, Burzette R, Stokes J. Differences in AIDS risk behavior among black and white bisexual men. Abstract presented at the Seventh International Conference on AIDS; June 1991; Florence, Italy.

Knowledge and Attitudes about AIDS among Corporate and Public Service Employees

Judith K. Barr, ScD, Joan M. Waring, MA, and Leon J. Warshaw, MD

ABSTRACT

Background. We examined the relationship between workplace AIDS education efforts and workers' knowledge about HIV transmission and their attitudes toward coworkers with AIDS.

Methods. Questionnaires were mailed to corporate and public service workers at 12 work sites to ascertain the extent of their knowledge about AIDS and their attitudes toward coworkers with AIDS. Each work site had offered an AIDS education program. The average response rate was 40%; 3460 workers returned questionnaires.

Results. Respondents' knowledge was largely consistent with available scientific evidence. However, a substantial minority still believe HIV infection can be transmitted through casual contact. Over 30% endorse the screening of new employees for AIDS, and 23% would fear contagion from an infected coworker. Thirty percent of the respondents expressed skepticism about the veracity of information from government sources and the scientific community. Work site comparisons show that where educational programs are minimal, employees know less about HIV transmission and hold more negative attitudes.

Conclusion. Comprehensive workplace AIDS education programs can reinforce workers' knowledge about HIV transmission, thereby fostering more favorable views toward coworkers with AIDS. (*Am J Public Health.* 1992;82:225–228)

Introduction

National surveys[1] and public opinion polls[2–4] have shown that a sizeable portion of the adult population is misinformed about HIV/AIDS. Negative attitudes toward people with AIDS (PWAs) have been associated with a lack of knowledge about HIV/AIDS among the general public.[5] Studies report an association between strong workplace HIV/AIDS education efforts[6,7] and positive attitudes about HIV/AIDS in the workplace. Workplace programs increased knowledge about HIV/AIDS transmission among public high school teachers and supervisors[8] and among employees in businesses.[9] University employees' attendance at an HIV/AIDS education program was associated with positive attitudes about HIV/AIDS.[10]

This paper examines relationships between employees' knowledge about HIV transmission and their attitudes toward coworkers with HIV/AIDS. Because HIV/AIDS education in the workplace is an important source of information for workers,[11] ascertaining what employees know and think can be useful in designing programs responsive to their needs and concerns.

Methods

Employees at eight corporate work sites and four public agencies were surveyed about their knowledge and attitudes about HIV/AIDS. Each of the participating organizations had implemented an HIV/AIDS education program, ranging from simple distribution of literature to extensive efforts including printed materials, video or film presentations, and discussion sessions with medical experts; some programs were compulsory.

The sample work sites were selected because of differences in approaches to HIV/AIDS education, the diversity of work organizations and work forces, and management's interest in and willingness to cooperate with the study. The work sites were dispersed geographically (New York state, mid-Atlantic, mid-West) and included a national sales force; they represented different industries (pharmaceutical, telecommunications, banking) in the private sector and agencies in the public sector.

The questionnaires, accompanied by a cover letter from a company or agency official, were mailed in the fall of 1988. Respondents included 2836 corporate employees and 624 public service workers (n = 3460), representing an average 40% response rate (Table 1). Although the possibility of nonresponse bias cannot be ruled out, respondents and the total work force at each site did not differ appreciably in proportions of women to men (gender was the only characteristic available for comparison).

The self-administered questionnaire was adapted from the National Health Interview Survey.[12] A transmission knowledge score was formed from 12 items about the likelihood of getting HIV/AIDS through casual contact (Cronbach's alpha = .91). A workplace attitude score was formed from five items assessing at-

Judith K. Barr and Leon J. Warshaw are with the New York Business Group on Health, Inc. Joan M. Waring is with The Equitable, New York, NY.

Requests for reprints should be sent to Judith K. Barr, ScD, Associate Director for Programs, New York Business Group on Health, Inc, 622 Third Avenue, 34th floor, New York, NY 10017-6763.

This paper was submitted to the journal November 8, 1990, and accepted with revisions June 10, 1991.

TABLE 1—Response Rates by Industry

	No. Sent[a]	Rate, %	No. Respondents
Private industry (total)	6728	42.2	2836
Pharmaceutical	5278	38.2	2018
Telecommunications	800	48.5	388
Banking	650	66.2	430
Public agencies (total)	2000	31.2	624
Uniformed	1500	29.0	435
Nonuniformed	500	37.8	189
Total	8728	39.6	3460

[a]Questionnaires and self-addressed postpaid envelopes were mailed by each participating organization. Respondents returned the questionnaires directly to the New York Business Group on Health. There was no identifying information on the questionnaire; a mark on the envelope identified the site.

TABLE 2—Comparison of New York Business Group on Health (NYBGH) Study Respondents (n = 3460) with Sample from National Health Interview Survey (NHIS)[12] on Selected Items: Knowledge about Transmission Through Casual Contact

	NHIS, %		NYBGH, %	
	Incorrect[a]	Correct[b]	Incorrect[a]	Correct[b]
What is the likelihood of getting AIDS or the AIDS virus from				
Living near a hospital/home for AIDS patients	12	88	4	96
Attending school with child who has AIDS virus	16	84	5	95
Eating in restaurant where cook has AIDS virus	38	62	17	83
Shaking hands/touching someone with AIDS virus	15	85	3	97
Working near someone who has AIDS virus	20	81	4	94
Sharing eating utensils with someone who has AIDS virus	41	59	20	80
Being coughed/sneezed on by someone with AIDS virus	40	60	21	79
Kissing cheek of person with AIDS virus	15	85	4	95
Kissing (with exchange of saliva) a person with AIDS virus	63	36	44	56
Using public toilets	29	71	9	91
Mosquitos or other insects	47	53	21	79
Donating or giving blood	—	—	43	58

Note. Percentages are rounded and may not add to 100.
[a]Incorrect responses included "very/somewhat likely" and "don't know."
[b]Correct responses included "very/somewhat unlikely" and "not possible."

titudes toward coworkers with AIDS (Cronbach's alpha = .85).

Nearly one fifth (18%) of the respondents (n = 627) wrote comments in the space provided on the questionnaire. Content analysis yielded 814 meaning units, which were classified into nine substantive categories.

Results

Knowledge

Employees' knowledge about HIV transmission was generally consistent with available scientific data (Table 2). Respondents from three sites (two public and one private) with minimal workplace programs (e.g., literature only, "one-shot" program, voluntary rather than compulsory) had, on average, knowledge scores significantly lower than those of respondents from the sites with more comprehensive programs (Duncan Multiple Range Test; $P < .05$).

Although these employees were more knowledgeable than a national sample[12] surveyed at about the same time (Table 2), sizable numbers still believed that HIV/AIDS can be transmitted through casual contact. Fifteen percent said transmission is possible through shared eating utensils; 17% said HIV/AIDS can be transmitted through coughing or sneezing. Thirty-nine percent thought it very or somewhat likely that HIV/AIDS can be spread through saliva in kissing.

A major theme of the volunteered written comments was concern about the risk of HIV infection through casual contact.[13] Thirty percent of the respondents expressed skepticism about the accuracy of public information, because scientific findings are sometimes reversed; lack of trust in scientific authorities; and worry that some messages are simply attempts to calm public fears. For example:

> "I feel that although we are studying the disease we still don't know everything. Ten years down the line I don't want to hear you have found AIDS *can* be transmitted by shaking hands."

> "I believe the AIDS problem is much more serious and contagious than the general public is led to believe, and that there is more known about AIDS than what is released to the public."

Attitudes

Negative and potentially discriminatory attitudes toward coworkers with HIV/AIDS were expressed by some respondents (Table 3). Nearly three fourths agreed that PWAs should be treated like other employees. However, a substantial minority said they would be uncomfortable eating lunch with PWAs, and nearly one in four would be "afraid of getting AIDS" from working near PWAs. Although few considered HIV/AIDS an acceptable reason for termination, nearly one third said employers should not hire PWAs. Workers in the uniformed public agencies had significantly less favorable attitude scores than employees in the other worksites (Duncan Multiple Range Test; $P < .05$).

The attitudes expressed in volunteered comments are consistent with these results. Many respondents (16%) were concerned about being protected from coworkers with HIV/AIDS. Comments showed that 12% favored testing, firing, or "putting them on disability" to guard uninfected employees against contact. A few (6%) expressed egalitarian views about treating PWAs fairly and compassionately.

Relationship between Knowledge and Attitudes

There was a strong positive correlation ($r = .437$, $P < .001$) between trans-

mission knowledge scores and workplace attitude scores: those who were more knowledgeable also had more favorable attitudes. Contingency analysis shows that 22% of respondents both were most knowledgeable and had the most favorable attitudes (highest two fifths on both scores), and 24% both were least knowledgeable and had the least favorable attitudes (lowest two fifths on both scores).

Yet 12% of employees with a high transmission knowledge score had only moderately favorable workplace attitudes. Although very few (2%) respondents thought it was likely that HIV/AIDS could be contracted by working near someone with AIDS, a much larger proportion (23%) said they would be afraid of infection if they worked with a PWA. This discrepancy was found across all work sites (Figure 1).

Multiple linear regression analysis was used to examine whether other factors, besides knowledge about transmission, might predict attitude scores. The results indicate that transmission knowledge scores contributed most (15%) to variation in attitude scores. Gender, education, and self-assessed knowledge were also significant; women, the more highly educated, and those who believed they knew as much or more than others about HIV/AIDS had more favorable attitudes. These factors together explained 23% of the variance in attitude scores (R^2 = .234).

TABLE 3—Workers' Attitudes about People with AIDS (n = 3460)

Attitude Statement	% Who Agree
People with AIDS should be treated at work just like anyone else.	72.2
Employers should have the right to dismiss employees who have AIDS.	8.8
My employer should screen out prospective employees who have AIDS.	31.5
I would be uncomfortable eating lunch with someone who has AIDS.	30.4
I would be afraid of getting AIDS if I worked with someone who has AIDS.	23.4

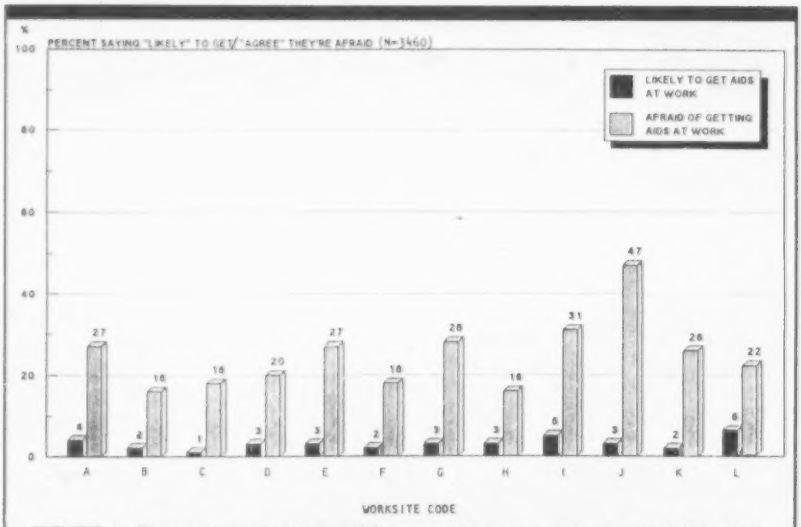


FIGURE 1—Perceptions of getting AIDS by working near someone with AIDS. Work sites are shown in descending order of employee participation in workplace AIDS education; corporate sites are A–F, K, and L; sites G–J are public agencies.

Discussion

The results show that employees having greater knowledge that HIV/AIDS is not transmitted through casual contact are less likely to hold negative views about coworkers with AIDS. Despite this strong relationship, extent of knowledge did not fully explain workplace attitudes. Our findings confirm observations that negative attitudes may persist even when people recognize their own low risk of becoming infected.[3] The volunteered comments suggest that many workers have heard the educational message but are not convinced of its accuracy; they are skeptical about HIV/AIDS information disseminated in the media and by medical experts.[10,13]

Comparisons of work sites in this study show that when employers' educational efforts are minimal, employees are less knowledgeable about HIV transmission through casual contact and less positive in their attitudes about HIV/AIDS in the workplace. Because employees find comprehensive workplace education programs highly informative and a credible source of HIV/AIDS information,[11] these results suggest that such programs should be used to reinforce the knowledge that casual contact is not a source of contagion. Negative and discriminatory attitudes can extend to coworkers who are perceived to be at risk for HIV or who are caring for PWAs.[14] Programs for HIV/AIDS education in the workplace can address these attitudes. By emphasizing the ways HIV/AIDS is *not* transmitted, these programs can reduce fears, thus fostering more tolerant attitudes toward coworkers who are themselves infected or who are caring for someone with HIV/AIDS.[15]

Workplace programs provide an important complement to public health HIV/AIDS education efforts. Private sector initiatives are under way to encourage and assist employers in providing education for employees. Efforts by community agencies and local health departments have made HIV/AIDS education available to workers through their employers. Continued collaboration between public health agencies and private organizations is essential to maintain the consistency, accuracy, and timeliness of information and to promote positive attitudes. At the same time, the efficacy of different types of workplace programs should be assessed. □

Acknowledgments

This research was supported in part by grants from the Robert Wood Johnson Foundation, the Health Services Improvement Fund, the NYNEX Corporation, and the United Hospital Fund of New York. Selected data were presented at the annual meeting of the American Public Health Association, New York City, October 3, 1990. The authors also thank the participating private employers and public agencies for their help in distributing the

questionnaires and the many employees for their time in completing them.

References

1. Dawson DA. AIDS knowledge and attitudes for July–September 1989: Provisional data from the National Health Interview Survey. No. 183. Public Health Service. Hyattsville, Md: National Center for Health Statistics; March 14, 1990. DHHS Pub. No. (PHS) 90-1250.
2. Singer E, Rogers TF, Glassman MB. Public opinion about AIDS before and after the 1988 US government public information campaign. *Public Opinion Quarterly*. 1991;55:161–179.
3. Blendon RJ, Donelan K. Discrimination against people with AIDS: the public's perspective. *N Eng J Med.* 1988;319:1022–1026.
4. Knowledge of AIDS is widespread; many taking preventive measures. *The Gallup Report.* February 1989;281:24–27.
5. Temoshok L, Sweet DM, Zich J. A three city comparison of the public's knowledge and attitudes about AIDS. *Psychol Health.* 1987;1:43–60.
6. Herold DM. AIDS in the workplace: what Georgia workers are thinking. Presented at conference on Managing AIDS in the Workplace, Georgia Institute of Technology, Atlanta, January 1989.
7. Herold DM. Employee attitudes toward co-workers with AIDS. Presented at meeting of the Academy of Management, Anaheim, Calif, August 1988.
8. Ruder AM, Flam R, Flatto D, Curran AS. AIDS education: evaluation of school and worksite based presentations. *NY State J Med.* 1990;90:129–133.
9. Feldmann TB, Bell RA, Purifoy FE, Stephenson JJ, Schwenhart A, Grisson S. Outcomes of AIDS training for supervisory and nonsupervisory personnel in the workplace. *J Occup Med.* 1990;32:103–109.
10. Hansen B, Booth W, Fawal HJ, Langner RW. Workers with AIDS: attitudes of fellow employees. *AAOHN J.* 1988;35:279–283.
11. Barr JK, Warshaw LJ, Waring J. Employee sources of AIDS information: the workplace as a promising educational setting. *J Occup Med.* 1991;33:143–147.
12. Hardy AM, Dawson DA. AIDS knowledge and attitudes for October and November 1988: Provisional data from the National Health Interview Survey. No. 167. Hyattsville, Md: National Center for Health Statistics; April 19, 1989. DHHS publication PHS 89–1250.
13. Barr JK, Warshaw LJ. *AIDS Education in the Workplace: What Employees Think.* New York, NY: New York Business Group On Health; 1990.
14. American Civil Liberties Union. *Epidemic of Fear: A Survey of AIDS Discrimination in the 1980s and Policy Recommendation for the 1990s.* New York, NY: American Civil Liberties Union; 1990.
15. Allard R. Beliefs about AIDS determinants of preventive practices and of support for coercive measures. *Am J Public Health.* 1989;79:448–452.

New Site for International AIDS Conference

The VIII International Conference on AIDS/III STD World Congress will now be held in Amsterdam, The Netherlands, July 19 to 24, 1992. Jointly sponsored by Harvard University and the Dutch Foundation AIDS Conference 1992, the Harvard-Amsterdam Conference will be cosponsored by the International AIDS Society and the World Health Organization. A new Call for Abstracts and Registration Brochure was issued in December 1991. The deadline for abstract submission and early registration is March 2, 1992.

For information on abstract submission and registration, write to the address nearest you: Harvard AIDS Institute, 8 Story Street, Cambridge, MA 02138, *or* VIII International Conference on AIDS/III STD World Congress, CLB, Plesmaniaan 125, Postbus 9190, 1006 AD Amsterdam, The Netherlands.

Workplace Compliance with a No-Smoking Law: A Randomized Community Intervention Trial

Nancy A. Rigotti, MD, David Bourne, MD, SM, Amy Rosen, PhD, John A. Locke, MPH, and Thomas C. Schelling, PhD

ABSTRACT

Background. Compliance with state and local laws restricting smoking in public places and workplaces has not been systematically evaluated.

Methods. We assessed workplace compliance with a comprehensive no-smoking law adopted in Brookline, Mass, and tested whether mailing information to businesses increased awareness of and compliance with the law. We conducted a random sample telephone survey of 299 businesses (87% response rate). Self-reported compliance was validated by direct observations.

Results. One year after its adoption, the law was popular with businesses. The prevalance of smoking restrictions, smoking policies, and no-smoking signs was 80%, 59%, and 40%, respectively. One third of businesses banned smoking. Full compliance with the law was low, however, because few businesses posted a copy of their smoking policy as required. The mailing increased employers' awareness of the law. Employers sent the mailing also reported better compliance, but this was not confirmed by direct observations.

Conclusions. The law was popular and contributed to a high prevalence of workplace smoking restrictions. Different interpretations of the law by policymakers and businesses seemed to explain why formal compliance was low. The mailing increased awareness of, but not compliance with, the law. (Am J Public Health. 1992;82:229–235

Introduction

Over the past two decades, states and communities have enacted laws restricting smoking in public places and workplaces at a rapid pace.[1] By 1988, 43 states and over 300 cities and counties had adopted a law limiting smoking in one or more public places.[2] The strictest laws require smoking restrictions at the workplace. As of 1989, laws in 14 states and 297 communities required workplaces to limit smoking.[3,4]

The intent of these laws, often termed clean indoor air laws, is to protect individuals from the discomfort and health hazards of involuntary smoke exposure. The laws may also decrease smoking prevalence by encouraging smokers to quit and discouraging nonsmokers from starting. Whether no-smoking laws will have these effects is not known because they have not been systematically assessed.[2,5] However, laws can be expected to have little impact on air quality or smoking behavior if individuals and businesses do not comply with them. To date, compliance has not been formally evaluated, and implementation has received little attention. Most laws delegate implementation to public health departments but do not fund this activity.[4] Health departments appear to regard the laws as self-enforcing, implying little need for surveillance because high compliance is expected.[2] Enforcement occurs in response to complaints, and a lack of complaints is considered evidence of good compliance and, presumably, reduced involuntary smoke exposure.[2,6] This assumption has not been tested, however.

The passage of a comprehensive no-smoking ordinance in Brookline, Mass, in 1987 provided the opportunity to evaluate workplace compliance with a community clean indoor air law and to test whether employers' awareness and compliance could be improved with intervention. We targeted businesses because the workplace is the primary site outside the home for passive smoke exposure.[7] Surveys indicate that 27% of US businesses and over one third of large corporations have smoking policies, but little is known about how smoking is handled in small businesses, which constitute the majority of American businesses.[1,2,8–11] Furthermore, little is known about how the business community responds to legislation requiring it to adopt smoking policies. The potential for a negative reaction has concerned some policymakers considering no-smoking laws.

We assessed workplace compliance with the Brookline no-smoking ordinance and conducted a randomized controlled trial of town businesses to determine whether a low-cost mailed program of education and surveillance could increase employers' awareness of and compliance with the law. The town's health department, which is responsible for implementing the law, conducted the intervention.

At the time of the study Nancy A. Rigotti was with the Department of Medicine, Massachusetts General Hospital, Harvard Medical School. David Bourne is with the Arkansas Health Department. Amy Rosen is with the Section of General Internal Medicine, Boston University Medical School. John A. Locke, is Director of Public Health, Brookline, MA. Thomas C. Schelling is with the Department of Economics, University of Maryland.

Requests for reprints should be sent to Nancy A. Rigotti, MD, General Internal Medicine Unit, Massachusetts General Hospital, Fruit St., Boston, MA 02114.

This paper was submitted to the journal February 12, 1991, and accepted with revisions June 24, 1991.

We hypothesized that employers sent information would be more aware of what the law required and, consequently, more likely to comply with it.

Methods

Brookline Bylaw

Brookline is a town of 55 062 residents in the Boston metropolitan area.[12] In 1982, there were 1455 businesses and government agencies in Brookline employing a total of 16 758 workers.[13] That year, Brookline became the third Massachusetts community to restrict smoking when the 250-member Town Meeting adopted a bylaw requiring restaurants with 40 seats or more to reserve 25% of those seats for nonsmokers (Bylaw XLI). Five years later, on June 1, 1987, the town meeting passed Bylaw XLII, which restricted or banned smoking in work sites and public places and enlarged restaurant nonsmoking sections to 50% of seats.

The worksite provisions of Bylaw XLII apply to businesses with three or more employees. The law requires employers to establish and post a work-site smoking policy. Employers must provide a nonsmoking working area for all employees. They are allowed, but not required, to designate smoking areas so long as the area does not expose nonsmoking employees or the public to smoke. No-smoking signs must be "conspicuously" displayed in nonsmoking areas. Upon request, a written copy of the policy must be provided to employees or to the director of public health.[14]

Workplace provisions of the bylaw became effective on March 1, 1988. Implementation and enforcement were the responsibility of the health department. The health director used local media, primarily the weekly town newspaper, to inform the public. Food service establishments were formally notified in January 1988 in a license renewal mailing. Inspectors subsequently checked for compliance during routine restaurant inspections. The health department sent no other information to citizens or businesses, except in response to complaints or questions. During the first year, the health department received 11 complaints about work-site smoking and informed those employers about the law. Work-site surveillance to monitor compliance was not done.

Intervention

We obtained a comprehensive mailing list of Brookline businesses from the Massachusetts Department of Employment and Training. The list included all companies using a Brookline address for payment of unemployment compensation taxes. All businesses with five or more employees (n = 535) were included in the trial. Businesses in the smallest category (less than five employees) were excluded because the bylaw applies only to work sites with three or more employees.

The 535 eligible businesses were randomly assigned to one of three groups. The health department mailed an information packet to companies in the two intervention groups in January 1989, 10 months after the law went into effect. No mailing was sent to controls. The packet sent to both intervention groups included (1) a copy of the bylaw, (2) a cover letter from the health director summarizing what the bylaw required of employers, and (3) a sample smoking policy. One intervention group also received a request from the health commissioner to return a written copy of its smoking policy. The estimated cost of preparing and mailing the information packet was $800.

Evaluation

We conducted a telephone survey of Brookline businesses in April 1989, 3 months after the mailing and 13 months after the law became effective. Firms were randomly selected from the 535 businesses in the trial. Of 459 firms telephoned, 53 (11%) could not be reached after four calls and 61 (13%) were not eligible. Interviews were completed in 299 of 345 eligible businesses, a response rate of 87%. Trained interviewers asked to speak to the owner, manager, or employee most familiar with the company's smoking rules. Respondents were informed that the survey was conducted for Harvard University and that data on individual firms were confidential.

The survey instrument, adapted from one we previously developed for a survey of Cambridge, Mass, businesses, contained 34 brief, closed-ended items and required 8 minutes to administer. Respondents were asked about their awareness and opinion of the law, their company's compliance, and their perception of the law's effect on employees, air quality, and costs. To understand how Brookline businesses handled smoking, we asked separate sets of questions about the company's smoking policy for employees, overall rules about smoking, smoking rules in six areas (private offices, shared work areas, public areas, conference rooms, hallways, cafeterias, and employee lounges), and posting of no-smoking signs. Type of business, work-force size, estimated employee smoking prevalence, and respondent's job title and smoking status were also recorded.

To confirm the validity of telephone self-reports, an observer made unannounced visits to all work sites reporting the presence of any one of the three observable outcome measures: a written smoking policy, posted smoking policy, or posted no-smoking sign. Additionally, the observer noted whether no-smoking signs were "conspicuously displayed" as required by law and whether employees or customers were smoking. Direct observations were made in 142 worksites, 48% of our sample. Smoking policies returned to the health department in response to the commissioner's request were reviewed by one rater, who judged whether each complied with the law.

This data collection permitted several separate assessments of compliance with the law: (1) respondents' self-rating; (2) respondents' reports of smoking rules, policies, and signs that constituted an operational definition of compliance; (3) direct observation of policies, signs, and smoking; and (4) analysis of written policies returned to the health department. For the purposes of the study, a business was in full compliance if it posted a written smoking policy at the work site, prohibited smoking in shared work areas and public areas, and posted signs in no-smoking areas.

Data Analysis

We compared the three groups on all endpoints, using chi-square and t tests to analyze the significance of differences in proportions and means.[15] Because the information sent to both intervention groups was very similar, we repeated the analysis, comparing a combined intervention group with controls to investigate the effect of a mailing. There were no major differences between the two intervention groups, and we report results using the combined intervention group. Multiple logistic regression was used to identify factors associated with awareness of the law, opinion of the law, and rules about smoking.[16]

Results

Characteristics of the Sample

Table 1 describes the 299 Brookline businesses surveyed. The average company was small; 73% had less than 20 employees. Nearly half were in service in-

dustries, and employee unions were rare. Employee smoking prevalence, estimated by respondents, was low, as was respondents' own smoking status. There were no significant differences between groups on any descriptive variable. The distribution of business size and type in the sample closely resembles the distribution for all Brookline businesses.[13]

TABLE 1—Characteristics of Brookline Businesses Surveyed[1]

	Information Only Group[a] (n = 94)		Information Plus Group[a] (n = 100)		Control (n = 105)		Total Sample (n = 299)	
	No.	%	No.	%	No.	%	No.	%
Size of work force								
1–9 employees	48	51	48	48	44	42	140	47
10–19 employees	17	18	28	28	33	31	78	26
≥ 20 employees	29	31	24	24	28	27	81	27
Type of business								
Service	46	49	47	47	48	45	141	47
Sales (retail or wholesale)	32	34	34	34	33	31	99	33
Finance, insurance, utility, real estate, transportation	10	11	11	11	12	11	33	11
Manufacturing, construction	6	6	8	8	12	11	26	9
Employee union								
Yes	7	8	4	4	2	2	13	4
Employees salaried								
< 25%	28	30	31	31	26	25	85	28
25–75%	21	22	23	23	26	25	70	23
> 75%	30	32	35	35	36	34	101	34
Unknown	15	16	11	11	17	16	43	14
Employee smoking prevalence[b]								
< 20%	56	63	68	71	63	66	187	67
≥ 20%	33	37	28	29	32	33	93	34
Respondent job title								
Owner/president/ VP	16	17	8	8	10	10	34	11
Director/supervisor/ manager	47	50	55	55	46	44	148	50
Assistant manager[c]	15	16	20	20	23	22	58	19
Other[d]	16	17	17	17	26	25	59	20
Smoking status								
Current smoker	21	23	22	22	22	21	65	22
Ex-smoker	26	31	33	33	33	31	92	31
Never smoker	46	49	44	44	50	48	140	47

Note. There were no statistically significant group differences on any variable.
[a]Both groups received mailed information about the law. The information plus group also received a request to submit their smoking policy to the health department.
[b]Estimated by respondent.
[c]Includes assistant manager, executive assistant, and administrative assistant.
[d]Includes treasurer and sales representative.

Awareness of the Law

Overall, 36% of businesses were aware of the town's no-smoking law. Companies sent the mailing were more likely to know about the law than were controls, suggesting that the mailing informed businesses (46% vs 18%, $P < .0001$; odds ratio [OR] = 3.88; Table 2). After adjusting, with multiple logistic regression, for potential confounders (business size and type, respondent job category and smoking status), the effect of group designation on awareness was unchanged (OR = 3.82; 95% confidence interval [CI] = 2.11, 6.93). No factor in Table 1 was independently related to awareness.

The 109 businesses aware of the Brookline bylaw were tested on their specific knowledge of six provisions highlighted in the mailing. Businesses correctly answered 64% of the questions, but the proportion of correct responses varied by question (Table 2). Businesses were best informed about the need to post no-smoking signs and ban smoking in public areas. They were least likely to know that providing a smoking area for employees was not required.

Opinion of the Law

One year after it took effect, the law was popular with businesses: 65% of respondents favored the law, 23% were neutral, and only 12% opposed it (Table 2). These proportions did not change when the analysis was limited to respondents aware of the law. The mailing did not influence attitudes about the law. On multivariate analysis, approval of the law was significantly higher in firms with nonsmoking respondents (OR = 4.22; 95% CI = 2.20, 8.17), work-site smoking restrictions (OR = 3.23; 95% CI = 2.23, 6.68), and fewer smoking employees (OR = 1.58; 95% CI = 1.12, 2.23). Approval was unrelated to company size or type.

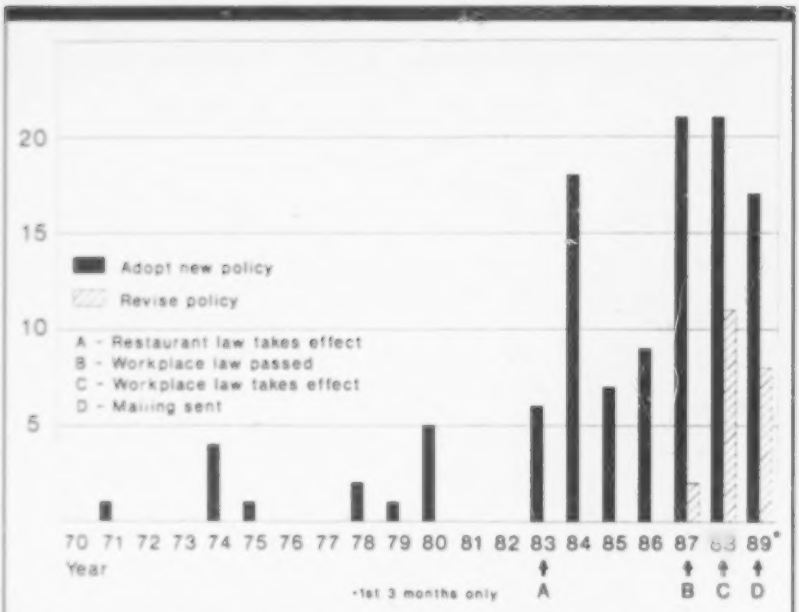
Compliance and Effects of the Law: Managers' Assessment

The 109 businesses aware of the law reported a high level of compliance and few problems. Ninety-five respondents (87%) considered their firm to be in compliance. Only three reported problems in complying; two problems involved employee opposition to smoking restrictions and one involved separating smokers and nonsmokers. Most companies reported that the law did not affect costs (n = 95, 87%) or air quality (n = 78, 72%); the remainder thought air quality had improved. Control and intervention groups did not differ on these questions.

The law also stimulated businesses to take action on smoking. According to respondents, the law prompted 39 firms aware of the law (36%) to adopt or revise a smoking policy. Forty (23%) of 177 companies with smoking policies cited the Brookline law as one reason their policy was adopted, and businesses sent the mailing cited the law more often than did controls (30% vs 10%, $P = .005$). Figure 1 illustrates the number of smoking policies adopted or revised each year among businesses whose managers recalled the date of policy adoption; 114 (64%) of 177 firms with policies could do so. Peaks of activity followed the adoption and implementation of no-smoking laws and the health department mailing. Two thirds (25/38) of firms implementing a policy after the law was adopted (1988 and 1989) were aware of the law.

Compliance: Self-Reports of Smoking Rules, Policies, and Signs

Table 3 displays measures of compliance with the law in all work sites and



Bars represent the number of smoking policies newly adopted (black bars) or revised (shaded bars) in each year, 1970–1989, among businesses who could recall the date of policy adoption. Arrows indicate when local smoking legislation was passed or implemented.

FIGURE 1—Smoking policies adopted or revised by Brookline businesses, 1970–1989.

distinguishes between smoking policies and smoking restrictions. Smoking restrictions were reported by 80% of work sites; 34% banned smoking and 46% restricted it to designated areas. Figure 2 presents the smoking rules for six specific areas. Over 65% of businesses banned smoking in shared work areas, public areas, conference rooms, and hallways. Smoking rules did not differ significantly between companies sent the mailing and controls, although those sent the mailing were more likely to ban smoking in shared work areas (70% vs 58%, $P = .068$).

Respondents were asked separately about their company's smoking policy. Fifty-nine percent replied that their company had a smoking policy for employees at that work site, but only half of these policies were written and only one third were posted at the workplace, as stipulated by law. Forty percent posted no-smoking signs. By our definition, 34 businesses (11%) reported full compliance with the law (Table 3). Companies sent the mailing were more likely than controls to report full compliance, a written smoking policy, and no-smoking signs, although the last difference was not statistically significant.

On multivariate analysis, smoking restrictions were more common in service businesses (OR = 3.48; 95% CI = 1.14, 6.13) and in firms whose top manager did not smoke (OR = 4.74; 95% CI = 2.09, 10.71). Smoking bans were more common in companies with a non-smoking manager (OR = 19.31; 95% CI = 2.56, 145.53) and fewer smoking employees (OR = 1.60; 95% CI = 1.12, 2.29). Restrictions were not independently associated with work-force size.

Compliance: Direct Observations

Direct observations in Brookline businesses confirmed self-reports of no-smoking signs in 98% of firms (119/121), although only 60% of signs were "conspicuously displayed" as required by law. In contrast, we validated few reports of written or posted smoking policies. We observed written policies in 6 of 89 companies reporting them (7%) and posted policies in 5 of 60 companies (8%). Employees often responded to a request for a written or posted smoking policy by pointing to a no-smoking sign.

Table 3 compares control and intervention groups on the validated endpoints of signs and policies. All written and posted policies were observed in businesses sent the mailing, but the number was so small that group differences were not statistically significant. By direct observation, only three Brookline companies (1%) fully complied with the law. Furthermore, a customer was seen smoking in 36% of businesses visited and a smoking employee observed in 14% (Table 3). Firms that banned smoking had significantly fewer smoking customers (15% vs 51%, $P < .0001$) and employees (6% vs 19%, $P = .037$).

Table 4 demonstrates that businesses aware of the law were more likely to comply with it.

Smoking Policies Returned to the Health Department

Half (n = 50) of businesses in the group asked to submit a written smoking policy did so. Nearly all used the demonstration policy provided by the health department. Twenty-nine policies (58%) were in compliance, indicating that smoking was restricted or banned and no-smoking signs were posted. The major reason for noncompliance was the misconception that firms without smoking employees need not have a smoking policy or post signs.

Discussion

This study assessed workplace compliance with a community clean indoor air law 1 year after the law's implementation. It also evaluated a low-cost intervention designed to increase employers' awareness and compliance with the law. Our discussion addresses these questions in order: (1) What was the prevalence of work-site smoking restrictions in Brookline 1 year after the law became effective? (2) Was the law responsible for these restrictions? (3) Did the intervention increase employers' awareness and compliance with the law? (4) What did town businesses think of the law?

1. One year after the law went into effect, the prevalence of work-site smoking restrictions in town businesses was high. Eight of ten firms had some smoking restrictions, one third banned smoking altogether, and two thirds banned smoking in shared work areas, where passive smoke exposure is likely to be highest. Two observations suggest that the situation was more complex, however.

First, there was considerable discrepancy among estimates of smoking policy prevalence. The prevalence of rules about smoking (80%) exceeded that of smoking policies (59%), and self-reported prevalence of smoking policies greatly exceeded that directly observed. Second, we observed customers and employees smoking in some businesses that reported not permitting it. These discrepancies could reflect employers' misrepresentation of their policy, failure to enforce their

policy, or misconception of what is a smoking policy and what the law requires businesses to do. The last two explanations are most likely. Smoking rules were probably not always enforced, and many businesses considered a no-smoking sign to be their written policy or felt no policy was needed because no employee smoked. Misrepresentation is less likely, given good agreement between self-reported and observed estimates of no-smoking-sign prevalence, a more concrete endpoint. It appears that the law was understood differently by policymakers who drafted it and businesses to whom it applied.

Because "smoking policy" was variously interpreted, we consider respondents' reports of "smoking rules" to be a better reflection of the actual prevalence of smoking restrictions in Brookline businesses. We conclude that the prevalence of smoking restrictions is high. Despite this, few Brookline firms complied fully with the law as written. Lack of awareness of the law and misinterpretation of the ordinance probably account for this finding. Nonetheless, it appears that the law achieved much of its intent—reducing exposure to passive smoke. The scarcity of complaints to the health department supports this interpretation but demonstrates that a lack of complaints does not guarantee compliance with a law. The frequency with which smoking was observed in businesses further indicates the limits of complaints as a measure of compliance.

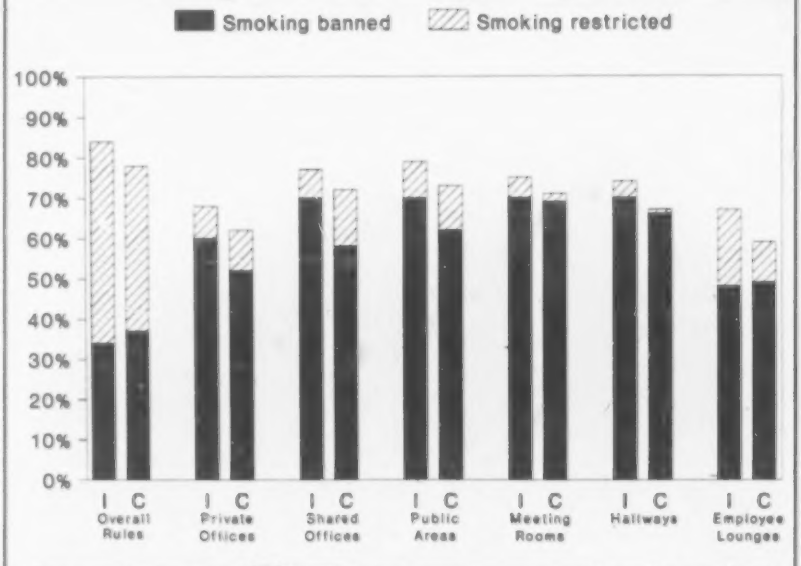
We suggest that differences in practice between large and small businesses account for some of employers' misconceptions about what the law required. While large businesses typically write formal policies to define acceptable employee behavior, small companies (as are most Brookline businesses) appear to make and enforce rules about smoking more informally and communicate them orally to employees, and they may not consider the rules to be policies.[17] Policymakers need to be aware of the practices of small businesses, which constitute the bulk of US companies, in order to draft work-site no-smoking laws that can be readily implemented by small and large firms.

2. To what extent was the Brookline law responsible for work-site smoking restrictions present 1 year later? The law was followed by a dramatic increase in the number of firms adopting smoking policies, and businesses aware of the law had more smoking policies. Companies aware of the law appeared to have been prompted to adopt smoking restrictions; however, without observations from a comparison community, we cannot exclude the possibility that smoking restrictions resulted from concurrent social change. Estimates of work-site smoking policy prevalence in other communities, with which to compare our data, are few. A 4% prevalence of written policies was found in New York City businesses in 1986, prior to a no-smoking ordinance.[18] In contrast, the prevalence of smoking policies and restrictions in Seattle manufacturing companies in 1988 was similar to Brookline, despite Seattle's lack of a work-site no-smoking law. However, Seattle bans smoking in public places and government work sites, and "public acceptance of smoking restrictions is widespread."[19] Although comparisons between surveys require caution, these data suggest that smoking policies may be less

TABLE 2—Employers' Awareness and Opinion of the Law

	Sent Mailing (n = 194)		Control (n = 105)		Total Sample (n = 299)	
	No.	%	No.	%	No.	%
Awareness of no-smoking bylaw	90	46 *	19	18	109	36
Opinion of the bylaw[a]						
Favor	127	65	68	65	195	65
Neutral	45	23	23	22	68	23
Oppose	22	11	14	13	36	12
Specific knowledge of law's requirements[b]	90		19		109	
Smoking area required (no)[c]	48	53	10	53	58	53
No-smoking area required (yes)	56	62	12	63	68	62
Ban smoking in public area (yes)	61	68	15	79	76	70
Post no-smoking signs (yes)	67	74	17	89	84	77
Post copy of policy (yes)	56	62	12	63	68	62
Provide copy of policy on request (yes)	57	63	12	63	67	61

[a]Results are similar when analyses are limited to those businesses aware of the bylaw.
[b]Questions asked only of those businesses aware of the bylaw.
[c]Correct answer in parentheses; columns show number of businesses answering question correctly.
*$P < .0001$.



Bars represent the percentage of Brookline businesses that ban or restrict smoking overall and in six specific areas at work. The figure compares the responses of companies sent the mailing (intervention group [I]) with controls (C).

FIGURE 2—Rules about smoking in Brookline businesses.

TABLE 3—Compliance with the Brookline Law

	Businesses Sent Mailing (n = 194)			Control Businesses (n = 105)		Total Sample (n = 299)	
	No.	%		No.	%	No.	%
Self-report (telephone interview)							
Rules about smoking							
Ban smoking	64	33		38	36	102	34
Restrict smoking	94	48		42	40	136	46
No restrictions	31	16		22	21	53	18
Unknown	5	3		3	3	8	3
Policy prevalence							
Any smoking policy	117	60		60	57	177	59
Written smoking policy	67	35	****	22	21	89	30
Posted smoking policy	45	23	**	15	14	60	20
No-smoking sign posted	86	44	*	35	33	121	40
Full Compliance[a]	28	14	***	6	6	34	11
Direct observation							
Policy prevalence							
Written smoking policy	6	3†		0		6	2
Posted smoking policy	5	3		0		5	2
No-smoking sign posted	83	43††		36	34	119	40
Full compliance[b]	3	2		0		3	1
Smoking observed[c]							
Employee smoking	14	15		4	12	18	14
Customer smoking	30	35		12	39	42	36

[a]By self-report, business has written, posted smoking policy and posted no-smoking signs and bans smoking in shared work areas and in public areas.
[b]Written, posted no-smoking policy and no-smoking sign observed and business bans smoking in shared work areas and in public areas.
[c]Limited to observed businesses with customer present (n = 116) or employee present (n = 126).
****P = .014; ***P = .023; **P = .066; *P = .094; †P = .007, Fisher's exact; ††P = .15

TABLE 4—Compliance in Businesses Aware and Unaware of the Law

	Aware (n = 109)		Unaware (n = 185)		
	No.	%	No.	%	P
Self-report (telephone interview)					
Rules about smoking					
Restrict smoking	92	84	144	78	NS
Policy prevalence					
Any smoking policy	78	72	99	54	.002
Written smoking policy	49	45	40	22	< .0001
Posted smoking policy	33	30	27	15	.001
No-smoking sign posted	56	51	65	35	.007
Direct observation					
Policy prevalence					
Written smoking policy	5	5	1	1	.028[a]
Posted smoking policy	4	4	1	1	.065[a]
No-smoking sign posted	58	53	61	33	< .0001

Note. NS = not significant.
[a]Fisher's exact test. All others are chi-square tests.

common in communities without laws, but considerable local variability exists. We conclude that smoking restrictions are more common in Brookline businesses than in many communities, due in part to the law.

3. What was the effect of the health department's mailing to businesses? The mailing increased awareness but not compliance with the law. It doubled the number of businesses aware of the law, but it informed fewer than half of town businesses about the law. Awareness is an important endpoint because compliance with a self-enforcing law can be expected only if those subject to it know about it. By itself, the mailing did not clearly increase compliance with the law, and the effect of the mailing was not enhanced by surveillance (e.g., requiring businesses to return a smoking policy to the health department). Surveillance might have been effective had the health department contacted noncompliant businesses it identified. Furthermore, analysis of returned policies provided information valuable for planning subsequent implementation programs. For the health department, major advantages of the intervention were its simplicity and low cost, but these factors also limited its effectiveness. Even with its modest effect, the mailing might be regarded as cost-effective; it cost $14.73 per business informed: $800/[(.46 − .18) × 194 businesses].

4. The law's popularity with employers 1 year after its adoption is an important additional finding. Employers aware of the law reported no adverse effects and did not perceive compliance to be burdensome. Not surprisingly, the law was more popular among respondents who did not smoke and in companies with smoking restrictions and few smoking employees. Managers' strong endorsement of the strictly worded Brookline statute is good news for policymakers in other communities considering no-smoking laws.

Methodologic strengths of this study include the following: we tested the intervention in a randomized controlled trial, assessed outcomes in a large random sample of town businesses, achieved a high response rate to a telephone survey, and validated self-reports by direct observation. Interpretation of study results must consider that this was a study of a single community. The Brookline sample resembles statewide and national distributions of business size, but it has significantly more service and fewer manufacturing businesses than comparison state and national samples.[10,20]

In summary, the Brookline no-smoking bylaw was popular with businesses, increased the adoption of work-site smoking policies, and contributed to a high prevalence of smoking restrictions in town businesses. Although full compliance with the law was low, largely because of misunderstandings in what the law required of businesses, the town's modest implementation effort achieved a good measure of the intent of the legislation. The health department's mailing increased companies' awareness but not compliance with the law. Further implementation efforts could improve awareness, compliance, and enforcement of the law. Nonetheless, the study demonstrates that a local health de-

partment can disseminate information to employers concerning a new requirement with a modest effort. □

Acknowledgments

This study was funded in part by Grant 88-6169 from the Henry J. Kaiser Family Foundation and a teaching and research schloarship awarded by the American College of Physicians to Dr. Rigotti.

This work was presented at the annual meeting of the American Public Health Association, October 1989.

We are indebted to Eva Jackson for assistance in data collection and analysis, to Pat Carney for manuscript preparation, and to David Hemenway, PhD, and Kenneth Warner, PhD, for reviewing the manuscript.

References

1. US Department of Health and Human Services. *The Health Consequences of Involuntary Smoking. A Report of the Surgeon General.* Washington, DC: US Department of Health and Human Services; 1986; DHHS publication no. (CDC) 87-8398.
2. US Department of Health and Human Services. *Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General.* Washington, DC: US Department of Health and Human Services; 1989; DHHS publication no. (CDC) 89-8411.
3. Tobacco-Free America. *State Legislated Actions on Tobacco Issues.* Washington, DC: Tobacco-Free America Legislative Clearinghouse; 1989.
4. Pertschuk M, Shopland DR. *Major Local Smoking Ordinances in the United States.* Washington, DC: Americans for Nonsmokers Rights and US Department of Health and Human Services; 1989; NIH publication no. 90-479.
5. Rigotti NA. Trends in the adoption of smoking restrictions in public places and worksites. *New York State J Med.* 1989;89:19–26.
6. Martin MJ. The San Francisco experience with regulation of smoking in the workplace: the first twelve months. *Am J Public Health.* 1986;76:585–586.
7. Repace JL, Lowrey AH. A quantitative estimate of nonsmokers' lung cancer risk from passive smoking. *Environment Int.* 1985;11:3–22.
8. Bureau of National Affairs. *Where There's Smoke: Problems and Policies Concerning Smoking in the Workplace. A BNA Special Report.* 2nd ed. Rockville, Md: Bureau of National Affairs; 1987.
9. Gottlieb NH, Hedl JJ, Eriksen MP, Chan F. Smoking policies among private employers and public agencies in Texas: a statewide analysis. *JNCI.* 1989;81:200–204.
10. Fielding JE, Piserchia PV. Frequency of worksite health promotion activities. *Am J Public Health.* 1989;79:16–20.
11. Dun & Bradstreet Corp. *Dun's Census of American Businesses 1989.* Parsippany, NY: Dun's Marketing Services Inc; 1989.
12. US Department of Commerce. *State and Metropolitan Area Data Book 1986.* Washington, DC: US Department of Commerce; Bureau of the Census; 1986.
13. Gibney FJ. *Brookline.* Boston, Mass: Massachusetts Department of Commerce, October 1984. Unpublished monograph.
14. Brookline, Mass, Bylaw Article XLII (November 16, 1988).
15. Wilkinson L. *Systat: the System for Statistics.* Evanston, Ill; Systat Inc; 1986.
16. SAS Institute. *SAS-PC Release 6.0.* Cary, NC; SAS Institute Inc; 1987.
17. Sorensen G, Rosen A, Pinney J, Rudolph J, Doyle N. Worksite smoking policies in small businesses. *J Occup Med.* In press.
18. Centers for Disease Control. Survey of worksite smoking policies—New York City. *MMWR.* 1987;36:177–179.
19. Kinne S. Prevalence and restrictiveness of smoking policies in King County, Washington, manufacturing worksites. *Am J Public Health.* 1990;80:1498–1500.
20. Horner ER, ed. *Massachusetts Municipal Profiles, 1988–89.* Palo Alto, Calif: Information Publications; 1989.

Attitudes Toward Genetic Testing of Amish, Mennonite, and Hutterite Families with Cystic Fibrosis

Shelley R. Miller, MS, and Robert H. Schwartz, MD

ABSTRACT

Background. We had the unique opportunity to study families with cystic fibrosis (CF) from an interrelated Amish–Mennonite kindred and from an isolated Hutterite kindred.

Methods. While obtaining blood for DNA linkage analysis to help find the CF gene, we asked family members to answer a questionnaire about attitudes and feelings toward genetic testing. The questionnaire was also administered to families receiving care at the University of Rochester. After DNA-based CF carrier testing was developed, we asked the families whether they wanted to know the results.

Results. Each cultural group felt differently about many issues. Between 31% and 78% approved of CF-carrier testing and 14% to 78% approved of prenatal diagnosis. The majority (59%) of Hutterite parents and siblings were opposed to carriers marrying carriers and most (82%) of the CF families wanted to know the results of carrier testing.

Conclusions. Cultural differences need to be understood before genetic services are offered. Also, most CF mutations and their clinical correlates (phenotypes) will need to be defined. There are as few as three CF mutations in the Hutterite population. It will be possible to offer them the option to choose accurate genetic counseling. It will be difficult to offer genetic services to other populations with more heterogeneous attitudes and many more identified and yet-to-be-identified CF mutations. (*Am J Public Health.* 1992;82:236–242)

Introduction

The cystic fibrosis (CF) gene is on the long arm of chromosome 7. Approximately 80% of the CF mutations at this locus have been identified.[1–3] Direct testing for these mutations and haplotype analysis of other linked DNA markers provide accurate prenatal diagnosis and carrier testing for all at-risk families in which one member has CF. When most of the mutations are identified, it will be possible to identify virtually all CF carriers whether or not they have close relatives with CF. Society must decide under what circumstances genetic testing is advisable and prudent for the general population.[4] Policy development for population screening needs to consider differences among cultural, ethnic, and racial groups.[5]

Between 1976 and 1988, we were introduced to members of an interrelated North American Amish–Mennonite CF kindred (23 related families with 41 CF-affected children) and to members of a genetically isolated Hutterite CF kindred (11 families with 19 CF-affected children). With the help of relatives and genealogists, we implemented a research project with three objectives.

Our first objective was to identify informative families and to obtain blood from volunteers to be used for DNA linkage analyses by collaborating laboratories searching for the CF gene.[6–9] Since these kindreds were from unique cultural backgrounds, our second objective was to try to understand their attitudes, intentions, and feelings regarding various CF-related issues. Later, when carrier testing in at-risk families became reliable and results of haplotype analyses were available, our third objective was implemented. We asked the Amish, Mennonite, and Hutterite CF families (parents of minors and adult siblings) whether they wanted to be informed of their carrier testing results.

We used a questionnaire survey method and compared the responses of Amish, Mennonite, and Hutterite families with a group of 95 CF families who were receiving care at the University of Rochester Medical Center. We also compared our results with those of a similar survey of 346 families in California by Kaback et al.[10] The Amish, Mennonite, and Hutterite responses offer a unique opportunity to better understand CF families from groups that are not well known to the medical world, have a high incidence of CF, and have distinct social views.

Methods

Subjects

The Amish are descendants of the Anabaptist faith of the 1600s.[11] They choose to live simply, in small farming communities, without modern conveniences. They believe in complete obedience to the Bible and its decisiveness and authoritativeness in all disputes.[12]

The Mennonites, also descendants of the Anabaptist movement,[11] have values similar to the Amish, but many are technologically modern and accept, and even promote, higher education.

Members of the Amish–Mennonite kindred are the direct descendants of a common ancestral couple, Christian J.

The authors are with the Department of Pediatrics, University of Rochester Medical Center, New York.

Requests for reprints should be sent to Robert H. Schwartz, MD, Allergy, Asthma, Immunology of Rochester, PC, 919 Westfall Road, Rochester, New York 14618.

This paper was submitted to the journal August 20, 1990 and accepted with revisions June 10, 1991.

(1766–1828) and Anna F. (1771–1816).[13] Their ethnic background is either Amish or Mennonite. They reside mostly in New York and Ontario, Canada; some Mennonites live in other states, including Alaska, Delaware, Indiana, Louisiana, Michigan, Ohio, and Pennsylvania. Their pedigree has been published.[7]

The Hutterites, also Anabaptists, are a genetic isolate of hard-working, mechanically modernized, religiously devout people who live in communal colonies on farm/ranch settings in Minnesota, North Dakota, South Dakota, Montana, and Washington, as well as in western Canada. The Hutterites have three distinct subgroups (Schmiedeleut, Dariusleut, and Lehrerleut) that originate from the three founding colonies.[14] Formal education is limited and birth control is not allowed. The Hutterites are extremely knowledgeable and productive farmers and animal breeders. Their own roots can be traced in accurate genealogies.[15] The Hutterite CF kindred has been published.[7,8] The parents of 11 CF sibships trace their ancestry to 62 of the 77 founders of the Hutterite population.[8]

Questionnaire

We modeled a questionnaire after one designed by Kaback et al.[10] It was reviewed and edited by a Canadian Mennonite genealogist sensitive to questions that might be offensive. The questionnaire elicited personal, religious, financial, and education data. It also contained questions about general CF knowledge, emotional impact and feelings regarding having a family member with CF, and related family planning issues. Attitudes about the possible use of carrier testing, prenatal diagnosis, and abortion were probed. Thoughts about possible effective treatment and a potential cure for CF were elicited.

TABLE 1—Questionnaire Recovery

	Sent	Returned Answered	Not Returned	Returned Not Answered	Moved, No Address
Amish					
Parents	10	7	2	1	0
Siblings	14	8	6	0	0
Aunt/uncle	28	14	11	1[a]	2
Subtotal	52	29	19	2	2
Mennonite					
Parents	31	19	11	1	0
Siblings	35	12	20	0	3
Aunt/uncle	171	98	64	3[a]	6
Cousins	22	22	0	0	0
Subtotal	259	151	95	4	9
Hutterite					
Parents	14	13	1	0	0
Siblings	18	16	2	0	0
Subtotal	32	29	3	0	0
Rochester					
Parents	146	65	75	0	6
Siblings	25	14	7	4	0
Subtotal	171	79	82	4	6
Intermarried					
Parents	1	1	0	0	0
Siblings	3	3	0	0	0
Aunt/uncle	17	15	0	0	2
Cousin	9	9	0	0	0
Subtotal	30	28	0	0	2
Total (%)	544	316 (58)	196 (36)	10 (2)	19 (4)
Kaback et al.[10]					
Parents	692[b]	362 (52)	46 (7)	126 (18)[c]	158 (23)

[a]Two of these individuals are deceased.
[b]346 households.
[c]Failed to meet study eligibility criteria (language, both parents together).

TABLE 2—Questionnaires Completed

	n	Male	Female	Parents	Siblings	Aunts	Uncles	Cousins
Amish	29	14	15	7	8	7	7	0
Mennonite	151	64	87	19	12	56	42	22
Hutterite	29	13	16	13	16	0	0	0
Rochester	79	34	45	65	14	0	0	0
Intermarried	28	16	12	1	3	5	10	9
Total	316	141	175	105	53	68	59	31
Kaback et al.[10]	414	203	211	414	...	...	...	...

TABLE 3—General Cystic Fibrosis Knowledge (% giving correct response)

	Parent: Recurrence Risk (1 in 4)			Unaffected Sibling: Carrier Risk (2 in 3)		
	Parents	Siblings	Total	Parents	Siblings	Total
Amish	43 (3/7)	12 (1/8)	27 (4/15)	0 (0/7)	0 (0/8)	0 (0/15)
Mennonite	89 (17/19)	83 (10/12)	87 (27/31)	10 (2/19)	33 (4/12)	19 (6/31)
Hutterite	54 (7/13)	87 (14/16)	72 (21/29)	0 (0/13)	19 (3/16)	10 (3/29)
Rochester	80 (52/65)	50 (7/14)	75 (59/79)	17 (11/65)	43 (6/14)	34 (27/79)
Intermarried	100 (1/1)	0 (0/3)	25 (1/4)	0 (0/1)	0 (0/3)	0 (0/4)
Total	76 (80/105)	60 (32/53)	71 (112/158)	12 (13/105)	25 (13/53)	16 (26/158)
Kaback et al.[10]	86 n = 362[a]	...	...	...	...	...

[a]184 mothers and 178 fathers

TABLE 4—Family Planning

	No. of Parents	Yes (%)	No (%)	Uncertain (%)	No Response (%)
Amish	7	0	14	72	14
Mennonite	19	53	47	0	0
Hutterite	13	61	15	15	8
Rochester	65	45	40	14	1
Intermarried	1	100	0	0	0
Kabeck et al.[10]	211[a]	48	40	11	...

Note. The question asked was, After having a child with CF and learning of your recurrence risk, did you change your planned family size?
[a]Mothers: 40 years of age or less (n = 184) and 41–45 years old (n = 27).

The questionnaire was mailed or hand delivered with an addressed, stamped return envelope. A cover letter asked for the respondents' cooperation by providing their confidential information to a research team working to understand CF. Each person was instructed to complete the questionnaire to the best of his or her ability and to call us collect or write us with any concerns, questions, or comments. The surveyed groups included 118 Amish–Mennonite and Hutterite parents and siblings; 273 Amish–Mennonite and Hutterite aunts, uncles, and cousins; and 171 Rochester parents and siblings. Each family had 1 or more living or deceased individuals who had been diagnosed with CF. Only those who were 18 years of age or older were asked to participate. The responses were anonymously coded and entered into a computer database.

When DNA analysis indicated carrier, noncarrier, or noninformative status of the siblings of CF individuals, we asked the adult siblings whether they wanted to know their carrier status and asked parents whether they wanted to know the status results of their minor children. We also asked the individuals whether they wanted us to inform their doctor of the results.

Results and Discussion

Response

Table 1 is a summary of questionnaire recovery. Table 2 is a summary of questionnaire response according to religious affiliation, sex, and relationship. Of 544 questionnaires, 316 (58%) were returned completed. Nineteen (3%) individuals had moved and left no forwarding address. A significant number of individuals did not return our questionnaire (199, 37%) or returned it unanswered (10, 2%). Possibly, we offended some people by asking personal questions. The length (67 questions) of the questionnaire may have been a deterrent. It is also possible that some individuals did not understand the instructions or questions and did not respond. Because the various religions have specific guidelines regarding such issues as abortion, an individual who differs from those guidelines might decline from responding on his or her own accord or on the advice of church elders.

Finances

We were unable to compare financial data since Hutterites practice a community sharing of goods, work, and life (The Chronicle of the Hutterian Brethren, 1987). When visiting families in the course of obtaining blood for DNA studies, it was our impression that the Amish were the poorest of our groups.

Education

The Hutterites (100%), followed by the Amish (90%), had the least formal education (i.e., less than 12 years of schooling). The majority of Rochesterians (94%), intermarrieds (93%), and Mennonites (72%) had completed high school, and 49% of intermarrieds, 47% of Mennonites, and 42% of Rochesterians had completed college.

Family Size

The average sibship size of an individual with CF was 6.2 in the Amish, 6 in the Hutterites, 3 in the Mennonites, and 2.7 in the Rochesterians. The largest family was found in the Hutterite population (17 children). Two Amish families had 10 children each. The largest Mennonite and Rochesterian families had 7 children each.

Religion/Ethnicity

Amish, Mennonites, and Hutterites are the three surviving Anabaptist groups sharing common beliefs about adult baptism, nonresistance, and simple living.

Members of the intermarried group were, by birth, Protestant, Catholic, or

TABLE 5—CF-Related Stress

	Parent Response (n = 105)						Sibling Response (n = 53)					
	No Stress		Slight–Moderate Stress		Much Stress		No Stress		Slight–Moderate Stress		Much Stress	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Amish	...		6	86	1	14	3	37	5	62	...	
Mennonite	...		10	53	9	47	2	17	10	83	...	
Hutterite	...		11	85	2	15	1	6	15	94	...	
Rochester	2	3	43	66	20	31	4	29	8	57	2	14
Intermarried	...		1	100	...		1	33	2	67	...	
Total	2	2	71	68	32	30	11	21	40	75	2	4

Note. CF = cystic fibrosis. The question asked was, How physically and emotionally demanding has caring for your child/children (sibling/siblings) been for you?

nondenominational; they married Amish or Mennonites but did not practice either of those religions.

The Rochester population consists primarily of those of Protestant (46%) and Catholic (43%) faiths; some are Jewish, some are Episcopalian, and some are Jehovah's Witnesses.

General CF Knowledge

Table 3 shows the percentages of individuals who correctly answered general CF questions regarding recurrence and carrier risks. Among parents and siblings, Mennonites (87%), Rochesterians (75%), and Hutterites (72%) were the most knowledgeable about recurrence risk to parents who have already had a child with CF. The Mennonite and Rochester parental responses closely resemble the responses among parents (85.9% correct) in the study of Kaback et al.[10] Those parents with the most formal education (Mennonites, Rochesterians, and intermarrieds) were the most knowledgeable. Except for the Amish, siblings were strikingly knowledgeable when asked this same question about their parents' recurrence risk.

Table 3 also indicates the percentage of correct responses regarding unaffected sibling carrier risk, a conceptually difficult question. It has been our experience that even some of the doctors and nurses who have helped us have incorrectly answered this question. Overall, 88% of parents and 75% of siblings answered incorrectly. Siblings were more knowledgeable than their parents when asked the unaffected sibling carrier risk. This is not surprising, since this statistic has more relevance for siblings. Again, those populations (Mennonites and Rochesterians) with more formal education were more knowledgeable.

Family Planning

Table 4 indicates that after having a child with CF and learning of their recurrence risk, the majority of Mennonite (53%) and Hutterite (61%) parents said they changed their planned family size, as did 45% of Rochesterians. The majority (72%) of Amish were uncertain about this matter.

Except for the Amish, our observations support the findings of Kaback et al.: having a child with CF and learning of the recurrence risk (correct or incorrect) has a major impact on future reproduction. Of those families who changed their planned family size, the majority (80% of Mennonites, 75% of Hutterites, and 100% of Rochesterians) had no more or fewer children than they would have liked. Although the numbers are small (n = 7), the uncertainty (72%) in the Amish is difficult to interpret. It may reflect a confusion with the question or a personal reflective uncertainty.

TABLE 6—Carrier Testing and Prenatal Diagnosis

		Carrier Testing for CF (%)			Prenatal Diagnosis for CF (%)		
	No.	Yes	No	Not Certain	Yes	No	Not Certain
Amish							
Parents	7	58	14	28	0	43	57
Siblings	8	37	50	12	0	25	75
Aunt/uncle	14	50	14	36	29	21	50
Total	29	48	24	28	14	28	59
Mennonite							
Parents	19	53	26	21	58	10	32
Siblings	12	67	8	25	50	25	25
Aunt/uncle	98	59	19	21	46	15	39
Cousin	22	64	9	27	86	9	5
Total	151	60	18	22	54	15	32
Hutterite							
Parents	13	54	15	31	31	15	54
Siblings	16	12	75	12	13	56	31
Total	29	31	48	21	21	38	41
Rochester							
Parents	65	78	11	11	80	8	12
Siblings	14	79	7	14	64	0	36
Total	79	78	10	11	77	6	16
Intermarried							
Parents	1	0	100	0	100	0	0
Siblings	3	67	0	33	67	33	0
Aunt/uncle	15	80	7	13	67	13	20
Cousins	9	45	22	33	100	0	0
Total	28	64	14	21	78	11	11
Kaback et al.[10]							
Parents	414[a]	...	...	...	82	8	10

Note. CF = cystic fibrosis. The questions asked were, Do you approve of carrier testing? and Do you approve of prenatal diagnosis for cystic fibrosis?
[a]211 mothers and 203 fathers.

TABLE 7—Wanted Communication of Results of Carrier Testing

	No. of Families	Yes (%)	No %	No Response
Amish	5	40	40	20
Mennonite	17	88[a]	0	12
Hutterite	11	82	9	9
Total	33	79	9	12

[a]One family wanted only their doctor to have the results.

CF-Related Stress

Table 5 indicates that the majority of each group of parents and siblings found their caretaking responsibilities to be slightly to moderately stressful. The majority of all of the groups found that caring for a child or sibling with CF had been slightly stressful. More siblings than parents found there was no stress, which probably indicates that parents are primarily responsible for child care.

Caring for children with CF can be physically and emotionally demanding. In our society, stress is a subjective factor that induces bodily or mental tension. It has both positive (motivation) and negative (fatigue, irritability) qualities. The most stressed groups were the Mennonite and Rochester parents. The Mennonites and Rochesterians, with their more modernized life-style and city dwelling, may have less emotional and physical support

TABLE 8—Carriers Marrying Carriers

	No.	Allow Carriers to Marry Carriers: Yes (%)	No (%)	Not Certain (%)
Amish				
Parents	7	14	0	86
Siblings	8	0	12	87
Aunt/uncle	14	29	36	36
Total	29	17	21	62
Mennonite				
Parents	19	42	5	53
Siblings	12	67	0	33
Aunt/uncle	98	25	25	49
Cousin	22	64	4	32
Total	151	36	18	46
Hutterite				
Parents	13	0	54	46
Siblings	16	0	62	37
Total	29	0	59	41
Rochester				
Parents	65	48	17	35
Siblings	14	64	7	29
Total	79	51	15	34
Intermarried				
Parents	1	0	100	0
Siblings	3	0	0	100
Aunt/uncle	15	33	20	47
Cousin	9	67	0	33
Total	28	39	14	46

Note. The question asked was, Do you think carriers should be allowed to marry carriers?

available to them and thus feel more stress than the traditional rural Amish and Hutterites.

Carrier Testing

Table 6 indicates that more than half of Amish (58%), Mennonite (53%), Hutterite (54%), and Rochester (78%) parents approved of carrier testing prior to its actual availability. Among siblings, the majority of Mennonites (67%) and Rochesterians (79%) approved, while the majority of Amish (62%) and Hutterites (87%) did not approve. The majority of aunts and uncles approved of CF-carrier testing.

The younger generation has mixed feelings. The older generation may be experiencing the well-known phenomenon of parents wanting their children to have a better life than they have had, including having healthy children.

Prenatal Diagnosis

Table 6 also indicates that the majority of Mennonites (54%), Rochesterians (77%), and intermarrieds (78%) approved of prenatal diagnosis for CF, while more Amish (86%) and Hutterites (79%) did not approve or were uncertain.

Rochesterians and the population of Kaback et al., most similar in life-style, were also most similarly aligned in their approval of prenatal diagnosis. The Amish and Mennonites, from the same religious roots, single themselves out as having the most unfavorable and uncertain reaction to prenatal diagnosis. Prenatal diagnosis may be more morally and religiously objectionable to this group.

Carrier Testing Results

Table 7 shows that the majority of Mennonites (88%) and Hutterites (82%) clearly wanted to know their carrier testing results. This may reflect their desire to have some control over pregnancy outcome—at the time of mate selection. Members of one Mennonite family wanted the results to be given only to their doctor.

The Amish were equally split between wanting to know (40%) and not wanting to know (40%), with one family providing no response. The sample size is too small to draw conclusions. Considering that some families wanted to know their results and considering their high degree of uncertainty as to whether carriers should be allowed to marry carriers, it is possible that some Amish families might use their carrier results to influence the mate selection of their carrier children.

Carriers Marrying Carriers

Table 8 shows that the majority of Hutterites (59%) did not think CF carriers should marry CF carriers. Rochesterians were the group most (51%) approving. This question generated no response or uncertainty in a majority of Amish and in almost half of the Mennonites and intermarrieds. The majority of Amish and Mennonite parents were uncertain about whether carriers should be allowed to marry carriers.

It appears that the Hutterite parents might use carrier testing information (Tables 6 and 7) to discourage carriers from marrying carriers. Perhaps their self-taught skills and knowledge of plant and animal reproduction carry naturally into the practice of their own reproductive lives and they are able to avoid certain marriages if they are at risk for having a child with a genetic disease. The Hutterites keep extensive records of their genealogy and know the relationship of each young courting couple. Rochesterians, approving of carrier testing and of carriers marrying carriers, might use their carrier knowledge to avail themselves of prenatal diagnosis. It is unclear whether the Amish and Mennonites would intervene in the marital selections of their children.

Abortion

Table 9 shows that Amish and Hutterite parents (100%) and siblings (100%) are opposed to or uncertain about the elective abortion of CF-affected fetuses. Mennonites were almost evenly split between approving (37%), disapproving (35%), and being uncertain (28%). More than half of Rochesterians (56%) and intermarrieds (71%) approved.

The Amish and Hutterites are most opposed to abortion under any circumstances. These two groups are the most socially, financially, and geographically bound to their church, colony, and community, and they are the least formally educated. The Mennonites, Rochesterians, and intermarrieds were more lenient in their views about abortion, with the Rochesterians more lenient than the Mennonites. These three groups favored elective abortion if a woman wants one and/or only in specific cases. A surprising finding was the break with tradition of a small group of Amish aunts and uncles who approved of elective abortion at a woman's request.

Subsequent Pregnancies

Table 10 indicates that our research supports the findings of Kaback et al. that

TABLE 9—Attitudes Toward Abortion

		Elective Abortion of CF-Affected Fetus			Should Be Able to Have an Abortion			
	No.	Yes (%)	No (%)	Not Certain (%)	If She Wants One (%)	In Specific Cases (%)	Never (%)	No Response (%)
Amish								
Parents	7	0	71	29	0	0	58	42
Siblings	8	0	38	62	0	12	75	12
Aunt/uncle	14	29	43	29	36	7	57	0
Total	29	14	48	38	17	7	62	14
Mennonite								
Parents	19	32	21	47	16	68	16	0
Siblings	12	42	42	16	25	67	8	0
Aunt/uncle	98	33	40	27	17	48	27	8
Cousin	22	59	23	18	55	27	9	9
Total	151	37	35	28	23	49	21	7
Hutterite								
Parents	13	0	61	38	0	23	62	15
Siblings	16	0	62	37	0	19	69	12
Total	29	0	62	38	0	21	65	14
Rochester								
Parents	65	57	32	11	38	43	17	2
Siblings	14	50	21	29	57	36	7	0
Total	79	56	30	14	42	42	15	1
Intermarried								
Parents	1	100	0	0	100	0	0	0
Siblings	3	33		33	33	100	0	0
Aunt/uncle	15	67	20	13	47	20	20	13
Cousin	9	89	0	11	78	22	0	0
Total	28	71	14	14	64	18	11	7
Kaback et al.[10]								
Parents	414	...	...	...	54	38	7.6	...

Note. CF = cystic fibrosis. The question asked was, Which statement best reflects your attitude? A. A woman should be able to have an abortion if she wants one. B. A woman should be able to have an abortion only in specific cases. C. A woman should not be able to have an abortion under any circumstances.

TABLE 10—Subsequent Pregnancies

		Families with Cystic Fibrosis in Last Child (%)
Amish	...	...
Mennonite	50	(6/12)
Hutterite	71	(5/7)
Rochester	73	(29/40)
Total	68	(40/59)
Kaback et al.[10]	65	(119/183)

Note. The question asked was, After the diagnosis of cystic fibrosis was made in your child, was that child the last child in your family? Parents responded as a couple.

after a CF diagnosis was made in a child, many families (68%) stopped having children and the child with CF tended to be the last child born in a sibship of two or more children. Conclusions about the Amish parents could not be made. Some of these families were uncertain about when the CF diagnosis was made and how it coincided with future pregnancies. In some cases, additional children had already been born before the diagnosis was made. One intermarried parent was counted in the Mennonite group with his Mennonite wife.

It is not surprising that the physical, emotional, and financial challenges of having a child with CF cause families not to risk having another child with CF and to stop having children, healthy or otherwise. This behavior occurred prior to the availability of reliable prenatal diagnosis.

Treatment

When asked whether they thought a totally effective treatment, capable of stopping all symptoms, would be found in the next 5 years, the Amish were the most optimistic and the Hutterites were the most pessimistic. The responses, in order of optimism, were as follows: Amish, 52% (n = 29); Rochesterians, 48% (n = 79); Mennonites, 44% (n = 151); intermarrieds, 39% (n = 28); and Hutterites, 24% (n = 29).

Cure

When asked whether a cure would be found within the next 5 years, Rochesterians were the most optimistic and Hutterites were again the most pessimistic. The responses, in order of optimism, were as follows: Rochesterians, 59% (n = 79); Mennonites, 54% (n = 151); Amish, 52% (n = 29); intermarrieds, 39% (n = 28); and Hutterites, 34% (n = 29).

Conclusion

The fury of CF research and discoveries has made genetic testing a valuable option to many at-risk families. When considering genetic testing of populations, it is important to identify differences in prevalences of various CF mutations within and between cultural, ethnic, and racial groups. Our collaborators and others have begun to report these differences.[9,16,17] In addition, any phenotypic differences associated with these mutations must be defined in order to provide accurate counseling.[18–20] Likewise, before genetic testing and counseling services are offered on a voluntary or routine basis, it is also important that cultural differences in beliefs, attitudes, feelings, and behavior be identified and understood.

In our study of Amish, Mennonite, Hutterite, and Rochester families with CF, 31% to 78% approved of CF-carrier testing; 14% to 78% approved of prenatal diagnosis. Being opposed to abortion and/or the abortion of CF-affected fetuses did not deter some from considering prenatal diagnosis. Each cultural group felt differently about many issues.

The majority (59%) of Hutterite CF parents and siblings were opposed to carriers marrying carriers, and most (82%) of the CF families wanted to know the results of carrier testing. By haplotype analysis, only three CF mutations have been found in the Hutterites.[9] One of these is the common ΔF508 deletion. When the other two mutations are defined, it will be possible to offer the Hutterite population accurate genetic counseling. For other populations, where many more identified and yet-to-be-identified CF mutations exist, accurate counseling and testing will be more difficult. □

Acknowledgments

This work was supported in part by National Institutes of Health Grant AM34917, Integrated

Genetics, and the Rochester Chapter of the Cystic Fibrosis Foundation.

We wish to thank Lorraine Roth for her help in genealogy and for her wisdom about her people—the Amish-Mennonites. Richard A. Doherty, MD, provided an empathic academic atmosphere in which this research was encouraged. Ken Morgan, PhD, and Mary Fujiwara, MS, our Canadian colleagues, were pivotal in helping us to understand and study the Canadian Hutterites. Ann Mc Mullen, RN, MS, provided liason with the Rochester CF families. We also wish to thank Jeff Shapiro for his journalistic guidance and enthusiasm and Janice Cosgriff for her weekends of dedication. Special thanks are also extended to Katherine Klinger, PhD, Paul Watkins, PhD, and Pat Stanislovitis, MS, who provided ongoing analyses of carrier status as new techniques evolved.

References

1. Kerem B, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene: genetic analysis. *Science.* 1989;245:1073–1080.
2. Riordan JR, Rommens JM, Kerem B, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science.* 1989;245:1066–1073.
3. Rommens JM, Iannuzzi MC, Kerem B, et al. Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science.* 1989;245:1059–1065.
4. Wilfond BS, Fost N. The cystic fibrosis gene: medical and social implications for heterozygote detection. *JAMA.* 1990; 263:2777–2783.
5. Cystic Fibrosis Genetic Analysis Consortium. Worldwide survey of the ΔF508 mutation—report from the Cystic Fibrosis Genetic Analysis Consortium. *Am J Hum Genet.* 1990;47:354–359.
6. Klinger K, Stanislovitis PO, Hoffman N, et al. Genetic heterogeneity of cystic fibrosis. *Nucleic Acids Res.* 1986;14:8681–8686.
7. Watkins PC, Schwartz R, Hoffman N, Stanislovitis P, Doherty R, Klinger K. A linkage study of cystic fibrosis in extended multigenerational pedigrees. *Am J Hum Genet.* 1986;39:737–743.
8. Fujiwara TM, Morgan K, Schwartz RH, et al. Genealogical analysis of cystic fibrosis families and chromosome 7q RFLP haplotypes in the Hutterite brethren. *Am J Hum Genet.* 1989;44:327–337.
9. Klinger K, Horn GT, Stanislovitis P, Schwartz RH, Fujiwara TM, Morgan K. Cystic fibrosis mutations in the Hutterite brethren. *Am J Hum Genet.* 1990;46:983–987.
10. Kaback M, Zippen D, Boyd P, Cantor R. Attitudes toward prenatal diagnosis of cystic fibrosis among parents of affected children. In: Lawson D, ed. *Cystic Fibrosis: Horizons.* New York, NY: John Wiley and Sons; 1984.
11. Hostetler JA. *Amish Society.* Baltimore, Md: Johns Hopkins University Press; 1980.
12. Naylor PR. *An Amish Family.* New York, NY: Lamplight Publishing Inc; 1974.
13. Roth L. *Various Genealogies.* Kitchener, Ontario, Canada: L. Roth; 1982.
14. Hostetler JA. *Hutterite Society.* Baltimore, Md: Johns Hopkins University Press; 1974.
15. Hofer AM, Hofer N, Gross DP, Gross JDM. *Hutterite Roots.* Freeman, SD: Pine Hill Press; 1985.
16. Rozen R, Schwartz RH, Hilman BC, et al. Cystic fibrosis mutations in North American populations of French ancestry: analysis of Quebec French-Canadian and Louisiana Acadian families. *Am J Hum Genet.* 1990;47:606–610.
17. Cutting GR, Antonarakis SE, Buetow KH, Kasch LM, Rosenstein BJ, Kazazian H. Analysis of DNA polymorphism haplotypes linked to the cystic fibrosis locus in North American Black and Caucasian families supports the existence of multiple mutations of the cystic fibrosis gene. *Am J Hum Genet.* 1989;44:307–318.
18. Kerem B, Buchanan JA, Durie P, et al. DNA marker haplotype association with pancreatic sufficiency in cystic fibrosis. *Am J Hum Genet.* 1989;44:827–834.
19. Campbell PW, Phillips JA, Krishnamani MRS, Maness KJ, Hazinski TA. Cystic fibrosis: relationship between clinical status and F508 deletion. *J Pediatr.* 199¹;118:239–241.
20. Beaudet AL. Invited editorial: carrier screening for cystic fibrosis. *Am J Hum Genet.* 1990;47:603–605.

The Accuracy of Medicare's Hospital Claims Data: Progress Has Been Made, but Problems Remain

ABSTRACT

Background. Health care databases provide a widely used source of data for health care research, but their accuracy remains uncertain. We analyzed data from the 1985 National DRG Validation Study, which carefully reabstracted and reassigned ICD-9-CM diagnosis and procedure codes from a national sample of 7050 medical records, to determine whether coding accuracy had improved since the Institute of Medicine studies of the 1970s and to assess the current coding accuracy of specific diagnoses and procedures.

Methods. We defined agreement as the proportion of all reabstracted records that had the same principal diagnosis or procedure coded on both the original (hospital) record and on the reabstracted record. We also evaluated coding accuracy in 1985 using the concepts of diagnostic test evaluation.

Results. Overall, the percentage of agreement between the principal diagnosis on the reabstracted record and the original hospital record, when analyzed at the third digit, improved from 73.2% in 1977 to 78.2% in 1985. However, analysis of the 1985 data demonstrated that the accuracy of diagnosis and procedure coding varies substantially across conditions.

Conclusions. Although some diagnoses and all major surgical procedures that we examined were accurately coded, the variability in the accuracy of diagnosis coding poses a problem that must be overcome if claims-based research is to achieve its full potential. (*Am J Public Health.* 1992;82:243–248)

Elliott S. Fisher, MD, MPH, Fredrick S. Whaley, PhD, W. Mark Krushat, MPH, David J. Malenka, MD, Craig Fleming, MD, John A. Baron, MD, MS, MSc, and David C. Hsia, JD, MD, MPH

Introduction

Administrative health care databases provide an increasingly accessible and widely used source of data for health care research.[1–10] However, the accuracy of hospital discharge data remains uncertain. Studies conducted in the 1970s found that the coding of nonclinical data, such as age, gender, and dates of admission and discharge, was highly accurate but that diagnoses and procedures were less reliably coded.[11–13] Recent studies continue to raise questions about the accuracy of diagnostic coding.[14–16]

The National Diagnosis Related Group (DRG) Validation Study, which provided the most comprehensive recent assessment of the accuracy of hospital discharge data, found an overall error rate of 20.8% in DRG assignment.[17] However, because many individual DRGs include diverse clinical conditions,[18] the findings could not be generalized to the coding of specific diagnoses or procedures. Also, the published results could not be directly compared with those of the Institute of Medicine (IOM) studies of 1977 and 1980, even though both the National DRG Validation Study and the IOM studies used similar methods. We reanalyzed the data from the National DRG Validation Study to address these issues.

Methods

Record Selection and Reabstraction

The methods of the National DRG Validation study have been described in detail elsewhere.[17,19] Briefly, the Office of the Inspector General, US Department of Health and Human Services, randomly sampled hospitals from each of three bed-size strata, excluding specialty hospitals and those in states not using prospective payment during the study period. In the second stage, a random sample of 30 Medicare discharges between October 1, 1984, and March 31, 1985, was selected from each of the 239 hospitals and copies of the medical records were requested. Of the 7076 discharges selected, 7050 charts (99.6%) were obtained. As previously reported, the sample accurately represented the population of all Medicare beneficiaries in prospective payment system jurisdictions.[17] Accredited records technicians, blinded to the coding in the original records, reviewed the medical records, selected the supportable diagnoses and procedures, and translated them into International Classification of Diseases, Ninth Edition, Clinical Modification (ICD-9-CM) numeric codes. Reviewers adhered strictly

Elliott S. Fisher, David J. Malenka, Craig Fleming, and John A. Baron are with the Department of Medicine and the Department of Community and Family Medicine, Dartmouth Medical School, Hanover, NH. Elliott S. Fisher and Craig Fleming are also with the Veterans Affairs Medical Center, White River Junction, Vt. Fredrick S. Whaley is with the Department of Community and Family Medicine, Dartmouth Medical School. W. Mark Krushat and David C. Hsia are with the Office of Analysis and Inspection, Office of the Inspector General, Department of Health and Human Services.

Requests for reprints should be sent to Elliott S. Fisher, MD, MPH, Center for Evaluative Clinical Sciences, Dartmouth Medical School, HB 7250, Strasenburgh Hall, Hanover, NH 03755-3862.

This paper was submitted to the journal November 26, 1990, and accepted with revisions June 10, 1991.

TABLE 1—Frequency with Which Reabstracted Principal Diagnoses Agree with the Computerized Abstract in the 1977 Institute of Medicine study of National Hospital Discharge Survey Data (IOM) and the 1985 Office of the Inspector General National DRG Validation Study (OIG)

Diagnosis (ICD-9-CM code)	Agreement of Hospital Abstract Principal Diagnosis: With IOM Reabstract (1977)	With OIG Reabstract (1985)	% Change, 1977 to 1985
Infectious and parasitic diseases (001-139)	62.6	65.4	2.8
Neoplasms (140–239)	57.4	84.3	26.9
Endocrine, nutritional, and metabolic diseases (240–279)	71.3	81.1	9.8
Diabetes (250)	65.7	88.2	22.5
Diseases of the blood and blood-forming organs (280–289)	58.2	64.7	6.5
Mental disorders (290-313)	80.8	69.6	−11.2
Diseases of the nervous system and sense organs (320–389)	85.7	91.4	5.7
Diseases of the circulatory system (390–349)	74.3	83.1	8.8
Cerebrovascular accident (438)	61.2	77.9	16.7
Congestive heat failure (428)	73.9	84.3	10.4
Diseases of the respiratory system (460–519)	73.6	65.8	−7.8
Acute bronchitis (466)	66.4	52.7	−13.7
Diseases of the digestive system (520–579)	81.0	81.7	0.7
Cholelithiasis (574)	85.8	89.0	3.2
Diseases of the genitourinary system (580–629)	78.1	83.7	5.6
Diseases of the skin and subcutaneous tissue (680–709)	73.0	78.5	5.5
Diseases of the musculoskeletal system (710–739)	71.3	82.5	11.2

Note. The table shows the frequency (weighted percent), within each diagnostic group, with which the individual *three-digit* code on the original computerized abstract agrees with the individual three-digit code on the reabstracted record. For instance, within "Neoplasms," if the original abstract listed as the principal diagnosis the code 162.2, "malignant neoplasm of the main bronchus," and the reabstracted record listed a primary diagnosis of 162.3, "malignant neoplasm of the upper lobe, bronchus or lung," this would be tabulated as showing agreement, even though the four-digit codes differ.

to ICD-9-CM and Uniform Hospital Discharge Data Set coding rules. The ICD-9-CM codes were entered into the GROUPER program to assign the proper DRG. A second blinded recoding of 5% of the sample by different medical records technicians, conducted to determine the reliability of the reabstraction process, revealed no significant discrepancies in DRG assignment (agreement = 0.95, $\kappa = 0.86$, $Z = 2.12$).

Comparison with Institute of Medicine Study

We chose to compare our data with those from the IOM study of the accuracy of the National Hospital Discharge Survey (NHDS) data[13] rather than the IOM study of Medicare data[12] because the sampling methods of the former were similar to those of the National DRG Validation study, and the 1977 NHDS discharge abstract was more similar to current discharge abstracts than were the Medicare abstracts of the 1970s. Although the earlier study used ICDA-8 rather than the currently employed ICD-9-CM, the broad diagnostic groupings we present are identical. To assess overall agreement, we determined the proportion of all reabstracted records that had the same principal diagnosis coded on the original record and on the reabstracted record at the specified level of truncation (third or fourth digit). We used the same method to determine overall agreement for procedures (except that truncation was at either the second or third digit). To assess agreement within NHDS diagnostic categories, we determined the proportion of reabstracted records with a specified principal diagnosis that had the same principal diagnosis coded on the original record. Calculations were weighted according to our sampling weights to ensure comparability with the IOM study.

Diagnosis- and Procedure-specific Analyses of Coding Accuracy

To analyze the accuracy of coding with respect to specific procedures and diagnostic conditions, three clinicians familiar with both health services research and ICD-9-CM coding rules aggregated codes into clinically distinct groups of diagnoses and procedures such as might be used to select patient cohorts for studies of the outcomes of medical hospitalizations or surgical treatments. Within each diagnosis and procedure group, we compared the codes as recorded on the computerized abstract submitted to the Health Care Financing Administration (original record) with the codes as determined by the Office of the Inspector General (OIG) reabstraction of the medical record (reabstracted record).

We evaluated the coding accuracy of the current data using the concepts of diagnostic test evaluation.[20] Within this framework, we consider the reabstracted record to be the "gold standard" for defining the presence of a diagnostic condition or the performance of a particular procedure. Thus, the sensitivity (true positive rate) is the conditional probability that a diagnosis (or procedure) within the specified group was coded on the original record given that a diagnosis (or procedure) within the specified group was actually present on the reabstracted record. Specificity (true negative rate) is the conditional probability that a diagnosis or procedure was not present on the original record, given that is was not coded on the reabstracted record. Positive predictive value is the conditional probability that a diagnosis (or procedure) was actually present on the reabstracted record, given that it had been coded on the original record. To calculate 95% confidence intervals for sensitivity and positive predictive value, we used an exact method based on the two-tail binomial test.[21] For the analyses presented in Tables 2 and 3 we calculated the point estimates using both weighted and unweighted data. Because the results were very similar and the use of the unweighted data allowed us to calculate confidence intervals, we present our findings based on the unweighted data.

TABLE 2—Prevalence and Accuracy of Diagnoses on the Computerized Abstract Prepared by the Fiscal Intermediary, Compared with the Office of the Inspector General (OIG) Reabstracted Data (n = 7050 cases)

Diagnosis (ICD-9-CM code)	Diagnosis in Any Position			Diagnosis in Principal Position		
	n[a] (%)	Sensitivity[b, d]	Predictive Value[c, d]	n[a] (%)	Sensitivity[b, d]	Predictive Value[c, d]
Colon cancer (153–154.8)	75 (1.1)	0.92 (.83, .97)	0.88 (.79, .95)	51 (0.7)	0.92 (.81, .98)	0.87 (.75, .95)
Lung cancer (162–162.9)	130 (1.8)	0.93 (.87, .97)	0.90 (.84, .95)	62 (0.9)	0.92 (.82, .97)	0.84 (.73, .92)
Breast cancer (174–174.9)	33 (0.5)	0.97 (.84, 1.00)	0.84 (.69, .94)	24 (0.3)	0.96 (.79, 1.00)	0.88 (70, .98)
Prostate cancer (185)	101 (1.4)	0.84 (.76, .91)	0.86 (.79, .93)	55 (0.8)	0.93 (.82, .98)	0.96 (.87, 1.00)
Diabetes (250–250.99)	902 (12.8)	0.84 (.81, .86)	0.85 (.83, .88)	140 (2.0)	0.86 (.80, .92)	0.86 (.79, .91)
Dementia (290–290.9, 331.0–331.2)	91 (1.3)	0.76 (.66, .84)	0.60 (.50, .69)	17 (0.2)	0.53 (.28, .77)	0.75 (.43, .95)
Hypertension (401–405.99)	841 (11.9)	0.82 (.79, .84)	0.63 (.60, .66)	59 (0.8)	0.83 (.71, .92)	0.72 (.60, .82)
Congestive heart failure (402.01, 402.11, 402.91, 428–428.9)	788 (11.2)	0.89 (.86, .91)	0.71 (.68, .74)	402 (5.7)	0.85 (.81, .89)	0.87 (.83, .90)
Acute myocardial infarction (410)	271 (3.8)	0.90 (.85, .93)	0.87 (.82, .91)	204 (2.9)	0.94 (.90, .97)	0.92 (.87, .95)
Unstable angina pectoris (411.1–411.8)	203 (2.9)	0.88 (.83, .92)	0.75 (.69, .80)	151 (2.1)	0.82 (.75, .88)	0.78 (.70, .84)
Cerebrovascular disease (431, 432.9, 434–434.9, 436)	226 (3.2)	0.92 (.88, .95)	0.71 (.65, .76)	190 (2.7)	0.94 (.89, .97)	0.81 (.75, .86)
Peripheral vascular disease (440–442.9, 433.9)	217 (3.1)	0.58 (.51, .64)	0.53 (.47, .60)	38 (0.5)	0.76 (.60, .89)	0.69 (.53, .82)
Chronic obstructive lung disease (491–494, 496)	1019 (14.5)	0.85 (.82, .87)	0.87 (.85, .89)	176 (2.5)	0.79 (72, .85)	0.68 (.61, .74)
Peptic ulcer disease (531–534.91)	137 (1.9)	0.92 (.86, .96)	0.75 (.68, .81)	76 (1.1)	0.92 (.84, .97)	0.89 (.79, .95)
Chronic renal failure (585, 586)	168 (2.4)	0.83 (.76, .88)	0.70 (.63, .76)	20 (0.3)	0.85 (.62, .97)	0.61 (.41, .78)
Rheumatoid arthritis (714.0)	65 (0.9)	0.83 (.72, .91)	0.81 (.69, .89)	20 (0.3)	0.85 (.62, .97)	0.85 (.62, .97)
Hip fracture (820–820.9)	125 (1.8)	0.96 (.91, .99)	0.94 (.89, .98)	119 (1.7)	0.97 (.92, .99)	0.96 (.91, .99)

[a]Count (%) with specified codes present on the reabstracted (OIG) record.
[b]Sensitivity is the proportion of all cases with a given code (or range of codes) present on the OIG reabstracted record that were found to have a code within the same range on the bills submitted to HCFA on behalf of the hospitals by the Fiscal Intermediary.
[c]Positive predictive value is the proportion of cases with a given code (or range of codes) present on the record submitted to HCFA by the Fiscal Intermediary that were confirmed on the OIG reabstracted records.
[d]The 95% confidence interval is given in parentheses.

Results

Comparison of 1977 with 1985

Overall, the percentage of agreement between the principal diagnosis on the reabstracted record and the original hospital record when codes were truncated at the third digit was 78.2% in 1985, compared with 73.2% in 1977. When codes were truncated at the fourth digit, we found 71.5% agreement, compared with 63.4% in the earlier study. For principal procedures, agreement at two digits was 76.2%, identical to the 76.2% reported in 1977, and agreement at three digits was 73.8% in 1985, compared with 71.4% in 1977. When grouped by principal diagnosis into the diagnostic categories used in the NHDS, the percentage of agreement between reabstracted records and original records varied widely in each study (Table 1). In all but two diagnostic groups (mental disorders and diseases of the respiratory system), the accuracy of the coding of principal diagnosis improved between 1977 and 1985, with the greatest percentage improvement occurring in the coding of neoplasms.

Accuracy of Diagnosis and Procedure Coding

Tables 2 and 3 present relevant measures of coding accuracy for selected diagnoses and procedures according to whether the condition or procedure was coded in the principal position or in any position on the discharge abstract. We present two measures of the accuracy of hospital coding: sensitivity, the proportion of cases with a given code present on the reabstracted record that were correctly coded by the hospital; and positive predictive value, the proportion of cases assigned a given code by the hospitals that the reabstraction process confirmed. Specificity is not presented because it was over 94% (generally over 99%) for all conditions and can be calculated from the data presented.

The accuracy of diagnosis coding varied widely (Table 2). When position on the discharge abstract was ignored, sensitivity ranged from a low of 0.58 for peripheral vascular disease to a high of 0.97 for breast cancer. The positive predictive value for the selected conditions also varied widely, from 0.53 for peripheral vas-

TABLE 3—Prevalence and Accuracy of Procedures on the Computerized Abstract Prepared by the Fiscal Intermediary, Compared with the Office of the Inspector General (OIG) Reabstracted Data (n = 7050 cases)

Procedure (ICD-9-CM codes)	Procedure in Any Position			Procedure in Principal Position		
	n[a] (%)	Sensitivity[b, d]	Predictive Value[c, d]	n[a] (%)	Sensitivity[b, d]	Predictive Value[c, d]
Cataract extraction (13.1–13.69)	84 (1.2)	0.96 (.90, .99)	1.00 (.96, 1.00)	80 (1.1)	0.98 (.91, 1.00)	0.98 (.91, 1.00)
Valve replacement (35.2–35.28)	9 (0.1)	0.89 (.52, 1.00)	1.00 (.63, 1.00)	9 (0.1)	0.78 (.40, 1.97)	1.00 (.59, 1.00)
Coronary artery bypass graft (36.1–36.19)	25 (0.4)	0.96 (.80, 1.00)	1.00 (.86, 1.00)	18 (0.3)	1.00 (.81, 1.00)	0.90 (.68, .99)
Cardiac catheterization (37.2–37.29, 88.5–88.56)	97 (1.4)	0.88 (.79, .93)	0.89 (.80, .94)	73 (1.0)	0.84 (.73, .91)	0.88 (.78, .95)
Pacemaker (37.70, 37.73–37.77)	23 (0.3)	0.91 (.72, .99)	1.00 (.84, 1.00)	19 (0.3)	0.89 (.67, .99)	1.00 (.80, 1.00)
Carotid endarterectomy (38.12)	29 (0.4)	0.90 (.73, .98)	1.00 (.87, 1.00)	29 (0.4)	0.90 (.73, .98)	1.00 (.87, 1.00)
Resection/repair abdominal aorta (38.14, 38.34, 38.44, 38.64)	14 (0.2)	1.00 (.77, 1.00)	0.93 (.68, 1.00)	14 (0.2)	0.93 (.66, 1.00)	0.93 (.66, 1.00)
Peripheral vascular surgery (38.18, 38.38, 38.48, 38.68, 39.25, 39.29)	34 (0.5)	0.94 (.80, .99)	0.97 (.84, 1.00)	29 (0.4)	0.83 (.64, .94)	0.92 (.75, .99)
Colectomy (45.7, 45.8)	49 (0.7)	1.00 (.93, 1.00)	0.98 (.95, 1.00)	43 (0.6)	0.98 (.88, 1.00)	0.89 (.77, .96)
Cholecystectomy (51.2–51.22)	77 (1.1)	0.99 (.93, 1.00)	(1.00) (.95, 1.00)	67 (1.0)	0.96 (.87, .99)	0.96 (.87, .99)
Hernia repair, inguinal or femoral (53.0–53.39)	85 (1.2)	0.99 (.94, 1.00)	0.98 (.92, 1.00)	80 (1.1)	0.98 (.91, 1.00)	0.98 (.91, 1.00)
Transurethral prostatectomy (60.2)	153 (2.2)	0.98 (.94, 1.00)	0.98 (.94, 1.00)	129 (1.8)	0.95 (.90, .98)	0.87 (.80, .92)
Hysterectomy (68.3–68.5)	30 (0.4)	1.00 (.88, 1.00)	0.97 (.83, 1.00)	27 (0.4)	1.00 (.87, 1.00)	0.96 (.82, 1.00)
Hip arthroplasty, total or partial (81.5–81.59, 81.6–81.69)	66 (0.9)	0.97 (.89, 1.00)	(1.00) (.94, 1.00)	65 (0.9)	0.95 (.87, .99)	0.98 (.91, 1.00)
Mastectomy (85.4–85.48)	24 (0.3)	1.00 (.86, 1.00)	0.94 (.86, 1.00)	22 (0.3)	0.95 (.77, 1.00)	0.95 (.77, 1.00)

[a]Count (%) with specified codes present on the reabstracted (OIG) record.
[b]Sensitivity is the proportion of all cases with a given code (or range of codes) present on the OIG reabstracted record that were found to have a code within the range on the bills submitted to HCFA on behalf of the hospitals by the Fiscal Intermediary.
[c]Positive predictive value is the proportion of cases with a given code (or range of codes) present on the record submitted to HCFA by the Fiscal Intermediary that were confirmed on the OIG reabstracted records.
[d]The 95% confidence interval is given in parentheses.

cular disease to 0.94 for hip fracture. A similar range in coding accuracy was found when the analysis was restricted to the principal diagnosis.

The coding of major procedures was substantially more accurate than the coding of most diagnoses (Table 3). Sensitivity, when position was ignored, ranged from a low of 0.88 for cardiac catheterization to over 0.95 for 10 of the 15 procedures. Positive predictive value, for principal procedures, ranged from a low of 0.89 for cardiac catheterization to over 0.99 for seven procedures.

Discussion

Recent published research on the accuracy of Medicare's hospital discharge data has focused on reimbursement issues and has reported findings in the aggregate rather than for specific clinical conditions.[17,22] The National DRG Validation Study reported 79.2% agreement in DRG assignment for discharges in 1985.[17] We found similar overall accuracy at the three-digit coding level for principal diagnoses in 1985. However, these aggregate results obscured substantial variations in the coding accuracy of individual conditions and procedures. Several medical conditions, such as hip fracture, acute myocardial infarction, and most cancers, were accurately coded, as were virtually all major surgical procedures. Many clinical conditions, however, were coded with limited accuracy.

Several limitations of our study must be kept in mind. First, relatively small numbers of records were available for use in assessing the coding of many individual conditions. Consequently, some of our estimates have wide confidence limits and it would be inappropriate to assume that the coding of conditions that we did not examine is either poor or excellent. Second, our study relied on data from the last 3 months of 1984 and the first 3 months of 1985, and recent work by Carter et al. suggests that coding accuracy has improved since 1985.[22] Reabstraction of a national sample of records by the SuperPRO revealed that the proportion of cases on which the hospital and SuperPRO agreed on the weight of the DRG improved from 77% in 1985 to 86% in 1988.[22] However, because aggregate results reveal little about the coding of individual conditions, and the relative accuracy of coding across individual conditions is unlikely to have changed substantially, our findings should prove useful even to researchers who plan to use only the more recent data.

The most important limitation of our study, however, is that the clinical accuracy of the coding cannot be directly in-

ferred from our findings. Like previous studies, the National DRG Validation Study was designed to test the accuracy with which hospitals adhered to coding rules,[23] rather than to assess the validity of the claims data as a source of information on patients' health status or medical treatments. Consequently, the findings we report do not directly answer the two major clinically relevant questions of whether a specific condition present *in the patient* was coded on the computerized abstract (sensitivity) and whether specific conditions coded on the abstract actually were present *in the patient* (predictive value positive). This distinction between coding conventions and clinical status may explain the low accuracy with which some medical conditions are coded. A diagnosis may have been present but not recorded on the computerized abstract either because there were five more serious conditions present or because the condition was judged not to contribute to the care provided to the patient.[23] If a condition is coded by one abstraction team but not the other, it seems likely that the condition was documented in the patient's chart but that the two abstractors disagreed about whether it should be recorded among the listed diagnoses. Such misclassification could lead to an underestimate of both the sensitivity and the positive predictive value of the coded diagnoses.

This distinction between coding rules and clinical status also provides the basis for a caveat about our finding that the coding of acute myocardial infarction was relatively accurate. Although their study was based on a small sample of hospitals and was conducted prior to the implementation of the prospective payment system, Iezzoni et al. found that 9% of cases coded as acute myocardial infarction at nonteaching hospitals and 42% of cases at teaching hospitals did not meet clinical criteria for this condition.[14] New rules have been implemented specifically to address the limitations of myocardial infarction coding, but the proper clinical interpretation of many coded diagnoses will remain uncertain until clear-cut clinical criteria for their use are established and promulgated. Surgical procedures, because of the precision with which they are defined and the accuracy with which they are coded, pose less difficulty in interpretation.

What are the implications of our findings? One common use of administrative databases is in research comparing outcomes across individual hospitals or among classes of hospitals.[7,24] Our demonstration of variability in condition-specific coding accuracy and previously reported differences across types of hospitals in aggregate coding accuracy[14,16,17] underline the importance of interpreting claims-based hospital comparisons with caution. However, the potential value of administrative health care data has been demonstrated in follow-up studies based on more complete and accurate data,[25,26] and it justifies the increasing support for continued, if cautious, pursuit of claims-based studies.[8]

A second major use of claims databases is in population-based studies of the outcomes of specific conditions or treatments.[8–10] The major advantages of such claims-based studies are their relative efficiency and their potential to avoid the possible selection bias inherent in institution-specific or other convenience samples. From a clinical perspective, such studies may offer valuable information about the short- and long-term outcomes of common treatments and conditions in the general population.[10,27–29] Our findings demonstrate that major surgical procedures and medical conditions that are accurately coded, such as hip fracture and cancers, can be reliably identified through the claims data. However, for clinicians to apply the findings of claims-based research to help their patients understand the likelihood of relevant outcomes of specific treatments, the study population must be stratified according to the full range of clinical characteristics that may influence the outcome of treatment, including age, gender, severity of illness, functional status, and comorbidity.[29] The current level of coding accuracy and diagnostic precision is clearly insufficient to achieve the level of detail that most clinicians would seek in such studies.

Further improvements in the precision and accuracy of hospital coding would facilitate health care research. However, achieving a high level of accuracy for all conditions will clearly require substantial effort and attention to the many causes of inaccurate coding that have long been recognized.[17,30–32] In the short term, two steps should be considered to help overcome the constraints imposed by the variable accuracy of diagnosis coding and the uncertain clinical implications of diagnostic codes. First, studies similar in scope to the National DRG Validation study should be repeated at regular intervals and should report not only coding accuracy—the degree to which hospitals adhere to Uniform Hospital Discharge Data Set coding rules—but also the clinical accuracy of the coding. Second, a mechanism should be established to facilitate investigators' access to medical records for cases identified through the claims data. Chart review would not only ensure proper case definition, but would also allow precise case-mix measurement and accurate description of the specific interventions received by each patient.

The Medicare administrative databases are a valuable resource collected at great expense. Our findings show that some conditions may currently be studied using the claims data alone and that modest progress has been achieved in improving the accuracy of diagnostic coding. In the long run, improving overall accuracy would enhance claims-based research. In the short run, the implementation of steps to supplement the currently available data would provide investigators with an even more powerful tool for use in evaluating the outcomes of hospital care. □

Acknowledgments

Support for this study was provided by the Agency for Health Care Policy Research (grant R18 HS05745) and the National Institute of Aging (grant R01.AG07146-01).

For their assistance with this project, the authors are indebted to Cathaleen A. Ahern, Barry L. Steeley, and Jane C. Tebbutt of the DHHS Office of Inspector General; Patricia E. Brooks, Anne B. Fagan, and Barton C. McCann of the US Health Care Financing Administration; Penni I. St. Hilaire of the US Public Health Service; Annette M. Delaney, Laurie H. Moore, Joseph S. Yarmus, and their staff at Baxter Health Data Institute; and John R. Boyle and Walter P. Weaver of the DHHS Office of the Assistant Secretary of Management and Budget. We would also like to thank John E. Wennberg, Wayne Ray, and an anonymous reviewer for their helpful comments. □

References

1. Farmer ME, White LR, Brody JA. Race and sex differences in hip fracture incidence. *Am J Public Health.* 1984;74:1374–1380.
2. Kellie SE, Brody JA. Sex-specific and race-specific hip fracture rates. *Am J Public Health.* 1990;80:326–328.
3. Fisher ES, Baron JA, Malenka DJ, et al. Hip fracture incidence and mortality in New England. *Epidemiol.* 1991;2:116–122.
4. Ray WA, Griffin MR, Schaffner W, Baugh DK, Melton JL. Psychotropic drug use and the risk of hip fracture. *N Engl J Med.* 1987;316:363–369.
5. McPherson K, Wennberg JE, Hovind OB, Clifford P. Small-area variations in the use of common surgical procedures: an international comparison of New England, England, and Norway. *N Engl J Med.* 1982;307:1310–1314.
6. Wennberg JE, Gittelsohn AM. Small area variations in health care delivery. *Science.* 1973;183:1102–1108.
7. Health Care Financing Administration.

Medicare Hospital Mortality Information: 1986. Washington, DC: US Dept of Health and Human Services; 1988.

8. Roper WL, Winkenwerder W, Hackbarth GW, Krakauer H. Effectiveness in health care: an initiative to evaluate and improve medical practice. *N Engl J Med.* 1988;319:1197–1202.
9. Roos NP, Wennberg JE, Malenka DJ, et al. Mortality and reoperation after open and transurethral resection of the prostate for benign prostatic hyperplasia. *N Engl J Med.* 1989;320:1120–1124.
10. Wennberg JE, Roos N, Sola L, Schori A, Jaffe R. Use of claims data systems to evaluate health care outcomes: mortality and reoperation following prostatectomy. *JAMA.* 1987;257:933–936.
11. Institute of Medicine. *Reliability of Hospital Discharge Abstracts.* Washington, DC: National Academy of Sciences; 1977.
12. Institute of Medicine. *Reliability of Medicare Hospital Discharge Records.* Washington, DC: National Academy of Sciences; 1977.
13. Institute of Medicine. *Reliability of National Hospital Discharge Survey Data.* Washington, DC: National Academy of Sciences; 1980.
14. Iezzoni LI, Burnside S, Sickles L, Moskowitz MA, Sawitz E, Levine PA. Coding of acute myocardial infarction: clinical and policy implications. *Ann Intern Med.* 1988;109:745–751.
15. Jencks SF, Williams DK, Kay TL. Assessing hospital-associated deaths from discharge data: the role of length of stay and comorbidities. *JAMA.* 1988;260:2240–2246.
16. Greenfield S, Aronow HU, Elashoff RM, Watanabe D. Flaws in mortality data: the hazards of ignoring comorbid disease. *JAMA.* 1988;260:2253–2255.
17. Hsia DC, Krushat M, Fagan AB, Tebbutt JA, Kusserow RP. Accuracy of diagnostic coding for medicare patients under the prospective-payment system. *N Engl J Med.* 1988;318:352–355.
18. Iezzoni LI, Moskowitz MA. Clinical overlap among medical diagnosis-related groups. *JAMA.* 1986;255:927–929.
19. Delaney AM, Hsia DC. *National DRG Validation Study: Final Report.* Waltham, Mass: Baxter Health Data Institute; 1987.
20. Sox HJ. Probability theory in the use of diagnostic tests: an introduction to critical study of the literature. *Ann Int Med.* 1986;104:60–66.
21. Conover WJ. *Practical Nonparametric Statistics.* 2nd ed. New York, NY: John Wiley & Sons Inc; 1986.
22. Carter GM, Newhouse JP, Relles DA. How much change in the case-mix index is DRG creep? Santa Monica, Calif: The RAND/UCLA/Harvard Center for Health Care Financing Policy Research; 1990. Publication R-3826-HCFA.
23. American Hospital Association. *ICD-9-CM Coding Handbook for Entry-Level Coders, with Answers.* Chicago, IL. 1979;58.
24. Dubois RW, Brook RH, Rogers WH. Adjusted hospital death rates: a potential screen for quality of medical care. *Am J Public Health.* 1987;77:1162–1167.
25. Dubois RW, Moxley JH, Draper D, Brook RH. Hospital inpatient mortality: is it a predictor of quality? *N Engl J Med.* 1987;317:1674–1680.
26. Maerki SC, Luft HS, Hunt SS. Selecting categories of patients for regionalization: implications of the relationship between volume and outcome. *Med Care.* 1986;24:148–158.
27. Fisher ES, Malenka DJ, Solomon NA, Bubolz TA, Whaley FS, Wennberg JE. Risk of carotid endarterectomy in the elderly. *Am J Public Health.* 1989;79:1617–1620.
28. Whittle J, Steinberg EP, Anderson GF, Herbert R. Use of Medicare claims data to evaluate the outcomes of elderly patients undergoing lung resection for lung cancer. *Chest.* 1991;100:729–734.
29. Fisher ES. Helping patients choose treatments: an important role for administrative health care databases. *Chest.* 1991; 100:595–596.
30. Demlo LK, Campbell PM. Improving hospital discharge data: lessons from the national hospital discharge survey. *Med Care.* 1981;19:1030–1040.
31. Lloyd SS, Rising JP. Physician and coding errors in patient records. *JAMA.* 1985;254: 1330–1336.
32. Cohen BE, Pokras R, Meads MS, Krushat WM. How will diagnosis related groups affect epidemiologic research? *Am J Epidemiol.* 1987;126:1–9.

Public Health Briefs

Cancer Anxiety and Attitudes toward Mammography among Screening Attenders, Nonattenders, and Women Never Invited

ABSTRACT

A mailed questionnaire survey was conducted among the following groups: 179 women who screened false positive at a free mammography screening; a random sample of 250 women who screened negative; 670 nonattenders of the screening; and a random population sample of 250 women who lived in another city and were not invited, but were otherwise comparable. The most frequently reported reason for nonattendance was not having the opportunity. Furthermore, only 18% of the nonattenders reported anxiety about breast cancer compared with 33% of the population sample ($P < .05$). Ninety-nine percent of the women who attended indicated a positive attitude toward mammography that had not been adversely affected by screening experiences. (*Am J Public Health.* 1992; 82:249–251)

Inger Torhild Gram, MD, and Suzanne E. Slenker, PhD

Introduction

Mammographic screening has been the only effective means of reducing breast cancer mortality.[1–5] However, several authors have questioned the magnitude of this mortality reduction and called attention to potential adverse effects of mammography screening.[6–11] The few available studies of this topic indicate that most women cope well with the screening situation and its consequences.[12–15] The purpose of this study was to investigate breast cancer anxiety and attitudes toward mammography among screening attenders, nonattenders, and women never invited to participate.

Methods

A free mammographic screening was offered to 4323 women aged 40 or older as part of the Third Tromsö Study conducted in Tromsö, Norway, in 1986 and 1987.[16] Altogether, 3653 (85%) accepted the mammogram. A total of 193 (5%) of the screenees required further evaluation, which for 40 subjects included a biopsy. Details of the screening and case-finding procedures are given elsewhere.[17]

Of the 193 women requiring further examination, only those 179 who were not diagnosed with breast cancer were eligible for the present study, and they constituted the false positive (FP) group. The three other groups in this study were a random sample of 250 women who screened negative (SN), the 670 nonattenders, and a random population sample (PS) of 250 women living in the nearby city of Harstad. The latter women were not invited to the screening but were otherwise comparable to the Tromsö women and thus served as the reference group.

A questionnaire concerning perceptions about mammography, frequency of breast self-examination, and anxiety about having breast cancer was designed, pilot tested, and then mailed to all study subjects in 1987 after the mammography screening was completed. A reminder questionnaire was sent out to all nonrespondents. Among nonattenders, 120 women (18%) were excluded from the study (8 had died, 17 had breast cancer, 32 had moved, and 63 were unknown at address). The response rate among the remaining women was 84% among the SN group ($n = 209$), 89% among the FP group ($n = 160$), 38% among the nonattenders ($n = 210$), and 66% among the PS group ($n = 164$). Subjects were classified as residing in rural areas if their travel distance to the mammography unit was about 30 minutes or more.

Statistical analyses of the data were performed using the Pearson chi-square statistic for categorical data and Student's t test for continuous data.[18] The analyses were performed using SAS programs.[19]

Results

The median age of the study population was 46 years (range of 40 to 61 years), and the mean years of schooling was 10. Risk factors for breast cancer—such as a family history, age at menarche, age at

Inger Torhild Gram is with the Institute of Community Medicine, University of Tromsö, Norway; Suzanne Slenker is with the Department of Social and Behavioral Sciences, Boston University, School of Public Health.

Requests for reprints should be sent to I. T. Gram, Institute of Community Medicine, University of Tromsö, Postuttak, N-9000 Tromsö, Norway.

This paper was submitted to the journal November 26, 1990, and accepted with revisions August 22, 1991.

TABLE 1—Prevalence (%) of Breast Cancer Anxiety at Follow-up and Recalled Prevalence of Anxiety 1 Year before, by Group: Tromsö, Norway, 1987

	Screening Negative (n = 209)[a]	False Positive (n = 160)[a]	Nonattenders (n = 178)[a]	Population Sample (n = 164)[a]
Follow-up	22*	40	18*	33
One year before[b]	28	38	13*	31

[a]Percentages are based on smaller numbers due to missing responses.
[b]Prior to the screening invitation for the Tromsö women.
*Significantly different ($P < .05$) when compared with the PS group during the same period.
**Significantly different ($P < .001$) when compared with the PS group during the same period.

TABLE 2—Proportion (%) of Women in the Screening Negative and False Positive Groups Describing the Screening Examination as Painful or Unpleasant: Tromsö, Norway, 1987

	Screening Negative (n = 205)	False Positive (n = 157)
Unpleasant only	18	26*
Painful only	4	4
Both	3	11*
Neither	75	59*

*Statistically significant, $P < .01$.

TABLE 3—Prevalence of Breast Cancer Anxiety at Follow-up among Women in the False Positive Group (n = 160) According to Selected Responses: Tromsö, Norway, 1987

Variables	Subjects (n)[a]	Prevalence (%)
Recalled anxiety 1 year before		
No	93	19**
Yes	58	72
Recalled anxiety at screening		
No	107	26**
Yes	41	76
Anxiety about potential workup examination		
No	79	20**
Yes	62	65
Information adequate in workup letter		
No	52	31*
Yes	81	48
Fear of having breast cancer at workup recommendation		
No	59	20**
Yes	76	54

[a]Totals do not add up to 160 due to missing responses.
*Statistically significant, $P < .05$.
**Statistically significant, $P < .001$.

first birth, and prior breast biopsy—did not vary by group. Data are not shown for these factors.

The nonattenders were more likely to live in rural areas than the attenders ($P < .001$); they were also more likely to be unemployed and never to practice breast self-examination than the PS group ($P < .05$). Thirty-two of the nonattenders had had a recent mammogram. This was considered a legitimate reason for nonattendance, and these women were removed from the analysis. The remaining women (n = 178) reported that not having the opportunity (39%); not wanting to participate in the Tromsö Study (15%); fear of X-rays (13%); concern about painful examination (4%); not receiving a personal invitation (4%); fear of discovering breast cancer (3%); and the potential of having a male examiner (3%), were the reasons for nonattendance. Some women gave more than one answer, 22% did not answer, while altogether 14% of the women claimed, without giving further explanations, that none of the listed factors was the rationale behind their nonattendance.

Table 1 shows that both the SN group and the nonattenders reported significantly lower breast cancer anxiety at follow-up than the PS group ($P < .05$). The nonattenders also recalled being less anxious about breast cancer 1 year before, compared with the PS group ($P < .001$). The changes within each group did not gain statistically significant P values.

Altogether, 84% of the women reported having been given adequate information in the screening invitation, and 79% reported the same about the screening examination. Among women receiving the workup letter and examination, 61% and 72% respectively, were satisfied with the information.

Table 2 shows that more women in the FP group than in the SN group experienced the screening examination either as unpleasant or as both painful and unpleasant ($P < .01$). However, the majority in both groups found it neither painful nor unpleasant.

Table 3 shows that, among the FP group, women who recalled having anxiety about breast cancer 1 year before (prior to the screening), anxiety about the anticipated workup examination, or fear of breast cancer upon receiving the workup recommendation were more likely to have breast cancer anxiety at follow-up, after the reassurance, than those who did not ($P < .001$). Women who were content with the information in the workup letter had a higher prevalence of breast cancer anxiety than those who reported the opposite ($P < .05$). No association was found between the prevalence of anxiety about breast cancer and how the information was perceived at the invitation, the screening, or the workup examination.

Ninety-two percent of the nonattenders and 99% of the attenders and the women never invited indicated willingness to participate in another free mammography screening in the future. Of the attenders, 99% said they would also recommend a similar screening to a friend.

Discussion

The present study shows that a high proportion of women in a general population, approximately one out of three, have anxiety about breast cancer. The results further suggest that having negative results on a screening mammogram decreases this prevalence and that women who elect not to attend a screening are less anxious about breast cancer than those who attend.

One strength of this study is that breast cancer anxiety among women who were invited to the mammographic screening can be compared with that of women who were not invited. Another is that reasons for nonattendance could be evaluated without taking the monetary cost of the mammogram into account.

One limitation of this study is the possibility of recall bias. Another is that the survey instrument was not of sufficient depth to explore the relationship of cancer anxiety to other related health-belief concerns. Nevertheless, our results, which suggest that anxiety about breast cancer may motivate attendance at breast cancer screening, are in accordance with other studies, which used survey instruments that focused on more attitude and belief dimensions—such as perceived susceptibility—than ours did.[20-22]

This study reveals a higher attendance rate than do most of the studies reviewed by Vernon et al.[23] The high acceptance may be due to the fact that the mammography screening was put in a broader context of a comprehensive health survey. Our results also reflect the fact that women living or working in the city center had easier access to the mammogram screening facility than those who did not. This inference of inconvenient locations as a significant factor in explaining nonattendance has been proposed in previous studies.[14,17,24]

We do consider the low response rate among nonattenders eligible for the study to be a limitation. The same problem was revealed in the study by Baines et al.[14] Although the 178 nonattenders may not be representative of all the women who declined, their answers should be of value in understanding reasons for nonattendance.

Our finding that 11% found the screening examination somewhat painful is in accordance with that of Baines et al.[14] but in contrast to that of Stomper et al., who found that only 1% reported the examination to be painful.[25] That more women in the FP group than in the SN group perceived the screening examination to be both painful and unpleasant may be because women in the FP group have breasts that are more difficult to examine due to size or density, thus necessitating a stronger and more painful compression of the breasts. These results indicate some drawbacks of screening that have also been revealed in other studies.[13]

The present study indicates that women who were anxious before the screening were more likely to remain so. Discouragingly, perceived adequate information does not seem to prevent anxiety about breast cancer among those who had to go through a workup examination. Additional measures need to be found to minimize this negative effect of the screening.

Nearly all the women taking part in the present study reported that they would attend another mammography screening and also recommend a screening to their friends. These results reflect a positive attitude toward mammography and a willingness to participate that has not been adversely affected by screening experiences. □

Acknowledgments

A summary of this paper was presented at the annual meeting of the American Public Health Association in New York City, September 30–October 4, 1990.

References

1. Shapiro S, Venet W, Strax P, Venet L, Roeser R. Ten- to fourteen-year effect of screening on breast cancer mortality. *J Natl Cancer Inst.* 1982;69:349–355.
2. Verbeek ALM, Hendriks JHCL, Holland R, Mravunac M, Sturmans F, Day NE. Reduction of breast cancer mortality through mass screening with modern mammography: first results of the Nijmegen Project, 1975–81. *Lancet.* 1984;1:1222–1224.
3. Palli D, Del Turco MR, Buiatti E, et al. A case-control study of the efficacy of a nonrandomized breast cancer screening program in Florence (Italy). *Int J Cancer.* 1986;38:501–504.
4. Tabar L, Fagerberg CJG, Gad A, et al. Reduction in mortality from breast cancer after mass screening with mammography. *Lancet.* 1985;1:829–832.
5. Tabar L, Fagerberg G, Duffy SW, Day NE. The Swedish two-county trial of mammographic screening for breast cancer: recent results and calculation of benefit. *J Epidemiol Community Health.* 1989;43:107–114.
6. Skrabanek P. False premises and false promises of breast cancer screening. *Lancet.* 1985;2:316–320.
7. Skrabanek P. The debate over mass mammography in Britain: the case against. *Br Med J.* 1988;297:971–972.
8. Wright CJ. Breast cancer screening: a different look at the evidence. *Surgery.* 1986;100:594–597.
9. Devitt JE. False alarms of breast cancer. *Lancet.* 1989;2:1257–1258.
10. Eddy DM, Hasselblad V, McGivney W, Hendee W. The value of mammography screening in women under age 50 years. *JAMA.* 1988;259:1512–1519.
11. Schmidt JG. The epidemiology of mass breast cancer screening: a plea for a valid measure of benefit. *J Clin Epidemiol.* 1990;43:215–225.
12. Dean C, Roberts MM, French K, Robinson S. Psychiatric morbidity after screening for breast cancer. *J Epidemiol Community Health.* 1986;40:71–75.
13. Ellman R, Angeli N, Christians A, Moss S, Chamberlain J, Maguire P. Psychiatric morbidity associated with screening for breast cancer. *Br J Cancer.* 1989;60:781–784.
14. Baines CJ, To T, Wall C. Women's attitudes to screening after participation in the national breast screening study. *Cancer.* 1990;65:1663–1669.
15. Gram IT, Lund E, Slenker SE. Quality of life following a false positive mammogram. *Br J Cancer.* 1990;62:1018–1022.
16. Jacobsen BK, Lund E. Level of education, use of oral contraceptives and reproductive factors: the Tromsö Study. *Int J Epidemiol.* 1990;19:967–970.
17. Gram IT, Lund-Larsen PG, Rosenlund AF, Störmer J. Mammografiscreening i Tromsö. Gjennomföring og resultat av den förste mammografiscreening i Norge. [Mammography screening in Tromsö: realization and results of the first mammography screening in Norway.] *Tidsskr Nor Laegeforen.* 1989;109:1040–1042. (In Norwegian)
18. Daniel WW. *Biostatistics. A Foundation for Analysis in the Health Sciences.* 4th ed. New York, NY: John Wiley & Sons; 1987.
19. SAS Institute Inc. *SAS/STAT™ Guide For Personal Computers.* 6th ed. Cary, NC: SAS Institute Inc; 1987.
20. Calnan M. The health belief model and participation in programmes for the early detection of breast cancer: a comparative analysis. *Soc Sci Med.* 1984;19:823–830.
21. Nielsen CC. Women's use of mammographic screening: the role of information, cues, and reinforcement. *Conn Med.* 1990;54:374–377.
22. Lerman C, Rimer B, Trock B, Balshem A, Engstrom PF. Factors associated with repeat adherence to breast cancer screening. *Prev. Med.* 1990;19:279–290.
23. Vernon SW, Laville EA, Jackson GL. Participation in breast screening programs: a review. *Soc Sci Med.* 1990;30:1107–1118.
24. Rimer BK, Keintz MK, Kessler HB, Engstrom PF, Rosan JR. Why women resist screening mammography: patient-related barriers. *Radiology.* 1989;172:243–246.
25. Stomper PC, Kopans DB, Sadowsky NL, et al. Is mammography painful? a multicenter patient survey. *Arch Intern Med.* 1988;148:521–524.

Medicaid Coverage of Screening Tests for Breast and Cervical Cancer

Leslie P. Boss, MPH, PhD, and Frederick H. Guckes

ABSTRACT

Although most women receive periodic Papanicolaou smear (Pap) for early diagnosis of cervical cancer, those who do not are more likely to be of lower socioeconomic status. Similarly, for the many women who do not receive periodic mammography for early diagnosis of breast cancer, cost has often been cited as a reason. Medicaid provides health benefits to roughly 9.4 million women of the appropriate ages for Pap tests and roughly 3.3 million women of the appropriate ages for mammography. The decision to provide such coverage is made on the state level. Of the 50 states and the District of Columbia, 49 provide some level of coverage for Pap smears, 39 for screening mammography. Knowing the extent of coverage allows public health professionals to take advantage of this funding source to provide services for lower-income women and may help initiate coverage in those states where it is not currently available. (*Am J Public Health.* 1992;82:252–253)

Introduction

Mammography coupled with physical breast examination and the Papanicolaou smear (Pap) are accepted screening tools for breast and cervical cancer, respectively, but not all women receive these tests at recommended intervals. A substantial barrier for routine screening is cost, and poor women are less likely to receive the tests at recommended intervals than are other women.[1]

The Medical Assistance Program, or Medicaid, mandates coverage of diagnostic and certain preventive services but not of breast and cervical cancer screening. Beyond services mandated by law, those covered by Medicaid are defined by each state based on a list of optional services. Medicare coverage, defined nationally, includes both screening mammography and Pap exams. Many elderly persons are eligible for coverage by both Medicaid and Medicare. When dual coverage exists, Medicare takes precedence.

Based on the per capita income of a state, federal Medicaid funds are available to match state expenditures. In 1989, the federal share ranged from 50% to 78%.[2] As benefits such as screening mammography are added to a state's Medicaid program, the state is responsible for payment and receives matching federal funds at the preassigned matching rate.

Knowledge of individual state practices regarding coverage of screening tests for breast and cervical cancer is important information for health practitioners and program planners. It can also be of value in efforts to mobilize support for increasing existing benefits.

Methods

A person responsible for cancer control activities in health agencies in each state (including the District of Columbia in this category) was asked about Medicaid coverage for screening for his or her state's population. When additional information was needed, the state Medicaid office was also contacted. General Medicaid statistics were obtained from the Health Care Financing Administration (HCFA). Because age-specific enrollment for females was not readily available for all states, the age distribution of all females in states reporting to CFA was applied to the total HCFA estimate of female Medicaid recipients to obtain estimates of the number of recipient females in the age groups of interest. The estimated numbers of screening-eligible women for whom Medicaid-financed screening coverage could be provided in each state were summed to estimate the total number of screens that could occur and be covered, given current Medicaid regulations.

Results

In 1989, there were approximately 24 million Medicaid recipients, 15 million (63%) of whom were female. Of all female recipients, approximately 9.4 million (64%) were aged 15 to 84 years, roughly corresponding to American Cancer Society guidelines for Pap smear screening. Similarly, roughly 3.3 million female Medicaid recipients (22%) were between the ages of 45 and 84, ages appropriate for routine mammographic screening.

Twenty-seven states allow coverage of Pap smears as a screening tool for all women eligible for Medicaid (Table 1). In another nine states and the District of Columbia, Pap smears are covered only with a physician's order. In five states, only the laboratory cost of the Pap smear is covered. In seven states, representing 22% of the total US population, women of childbearing age seen in family planning or maternity clinics are eligible for Medicaid coverage of Pap screening. Two states (2.8% of the US population) provide no coverage for screening for cervical cancer. It is estimated that there are 680 000 Medicaid eligible women for

Leslie Boss is in the Epidemiology Program Office at the Centers for Disease Control in Atlanta; Frederick Guckes is at Eastern Virginia Medical School.

Requests for reprints should be sent to Leslie P. Boss, Epidemiology Program Office, Centers for Disease Control, MS C08, 1600 Clifton Road NE, Atlanta, GA 30333.

This paper was submitted to the journal January 23, 1991, and accepted with revisions August 23, 1991.

whom even partial payment for a Pap exam is not covered.

In 23 states, coverage of screening mammography is provided for all Medicaid-eligible women; in an additional 15 states and the District of Columbia, coverage depends on a physician's order (Table 2). These 38 states represent 72% of the US population. No coverage is provided in 12 states. It is estimated that there are 570 000 Medicaid-eligible women ages 45 to 84 in these 12 states for whom mammography is not covered.

Legislation requiring third-party payors to provide coverage of screening for breast and cervical cancer exists in 33 and 8 states, respectively. The availability of Medicaid coverage in these states is shown in Tables 1 and 2.

TABLE 1—Medicaid Coverage of Pap Exams, by State,[a] October 1990

Extent of Coverage	States Providing	Legislation Yes[a]	No	Total
All eligible persons covered	AR, CA, CT, HI, IN, IA, KY, ME, MI, MT, NE, NY, NC, ND, OH, OR, TN, UT, VT, VA, WI, WY, **MA, MN, RI, SD, WV**[b]	5	22	27
Covered with physician's order	AK, AZ, DC, FL, ID, NH, NJ, SC, **DE, KS**	2	8	10
Only laboratory costs covered	LA, MD, MS, MO, OK		5	5
Prenatal, family planning services covered	AL, CO, IL, PA, TX, NM, WA		7	7
Not covered	GA, **NV**	1	1	2
Total number		8	43	51

[a]Includes the District of Columbia.
[b]States in bold print *have* legislation related to third-party payment for Pap exams.

TABLE 2—Medicaid Coverage of Screening Mammography, by State,[a] October 1990

Extent of Coverage	States Providing	Legislation Yes	No[a]	Total
All eligible persons covered	AR, CA, CT, IL, IA, KY, MA, MN, MO, NY, ND, OK, PA, RI, SD, TN, WA, WV, WI, IN, MT, OH, OR[b]	19	4	23
Covered with physician's order	AZ, CO, FL, KS, ME, NH, **AK, DE, DC, MS, NE, NJ, SC, UT, VT, WY**	6	10	16
Not covered	GA, HI, MD, MI, NV, NM, TX, VA, **AL, ID, LA, NC**	8	4	12
Total number		33	18	51

[a]Includes the District of Columbia.
[b]States in bold print do *not* have any legislation related to third-party payment for screening mammography.

Discussion

The current interest in and awareness of early detection of breast cancer has led to increased use of mammography, changes in payment practices, and legislative and regulatory activity at both the state and national levels.

Payment for diagnostic tests and appropriate follow-up care for cancer are generally covered by Medicaid, whereas screening procedures frequently are not adequately covered. However, expenditures for treatment of cancer can be decreased by cost-effective use of screening tests.[3,4] It is not known how many screening procedures have been described as diagnostic tests so that costs would be covered under current regulations.

Medicaid is one of many third-party payors of medical care. Third-party payment laws vary considerably in their mandate. In this assessment of coverage, we chose to include all existing legislation, given that such legislation signifies a certain level of consciousness in that legislative body concerning the importance of cancer screening.

In 1990, Congress passed Public Law 101–354, Preventive Health Measures with Respect to Breast and Cervical Cancers. A major part of its purpose is "to screen women for breast and cervical cancer as a preventive health measure," "to provide appropriate referrals for medical treatment of the women screened," and "to ensure, to the extent practicable, the provision of appropriate follow-up services."[5] The law is meant to provide a payment of last resort; it is not to pay for a service for which payment can be made by any other state or federal compensation or health benefits program such as Medicaid or Medicare or under any other insurance or prepaid program. The law requires that low-income women be given priority for these services. Unfortunately, implementation of this act will be gradual, as appropriations to support the law will be incremental.

Low-income women are particularly in need of opportunities for early diagnosis of cancer. Appropriate screening results in improved survival. Thus, legislators, caregivers, and public health professionals alike need to maximize the use of the benefits available to millions of American women through such government compensation programs as Medicaid. □

References

1. Dawson DA, Thompson GB. Breast cancer risk factors and screening: United States, 1987. *Vital Health Statistics;* Series 10, No. 172. Hyattsville, Md: National Center for Health Statistics; 1989.
2. Health Care Financing Administration. *A Statistical Report on Medical Care Provided by Title XIX of the Social Security Act.* Unpublished report.
3. US Congress, Office of Technology Assessment. *The Costs and Effectiveness of Screening for Cervical Cancer in Elderly Women.* Washington, DC: US Govt Printing Office; February 1990. Background paper OTA-BP-H-65.
4. US Congress, Office of Technology Assessment. *Breast Cancer Screening for Medicare Beneficiaries: Effectiveness, Cost to Medicare, and Medical Resources Required.* Washington, DC: US Govt Printing Office; November 1987.
5. Pub L No. 101–354, and Title XV amendment to the Public Health Service Act (42 USC §201 et seq).

Distribution of HIV Type 1 Infection in Childbearing Women in California

Frank J. Capell, BA, Duc J. Vugia, MD, MPH, Virginia L. Mordaunt, BA, William D. Marelich, MA, Michael S. Ascher, MD, Alan I. Trachtenberg, MD, MPH, George C. Cunningham, MD, Stephen S. Arnon, MD, and Kenneth W. Kizer, MD, MPH

ABSTRACT

The incidence of acquired immunodeficiency syndrome (AIDS) is increasing among California heterosexuals and children. To assess human immunodeficiency virus (HIV)-1 infection in childbearing women, we conducted a blinded serosurvey of newborns. Dried blood specimens taken from 99% of California births during the third quarter of 1988 (n = 135 808) and linked only to maternal demographic categories were tested for HIV-1 antibody by enzyme immunoassay and confirmed by Western blot. Period prevalence of HIV-1 infection was 7.4 per 10 000 childbearing women. Prevalence was highest for Black women and was also elevated for Hispanic and San Francisco Bay Area women. Findings suggest that California Hispanic women will make up an increasing proportion of new AIDS cases. (*Am J Public Health.* 1992;82:254–256)

Introduction

In California, acquired immunodeficiency syndrome (AIDS) and human immunodeficiency virus (HIV)-1 infection associated with heterosexual contact, injection drug use (IDU) among heterosexuals, and perinatal transmission are occurring with increasing frequency.[1] To assess HIV-1 infection levels in women of reproductive age, we conducted an unlinked survey of HIV-1 antibody prevalence in blood specimens taken from all newborns in California over a 3-month period. Seropositivity in such specimens indicates maternal HIV-1 infection and is associated with a 25% to 35% chance of infection in the newborn.[2] The survey was conducted as part of the Centers for Disease Control HIV-1 Family of Surveys.[3-5] In this report, we summarize HIV-1 infection levels in childbearing women classified by area of residence, race/ethnicity, and age. The present survey also provides relatively unbiased evidence on the extent of vertically acquired pediatric HIV-1 infection in California.

Methods

The study population consisted of all California women delivering live babies between July 1 and September 30, 1988, and screened for genetic diseases by the California Department of Health Services (CDHS) (accounting for approximately 99.8% of live births). A total of 136 861 births were eligible after excluding nonresidents and redundant records from multiple births.

The survey used newborn blood specimens and maternal demographic information, stripped of identifiers, originally collected for the CDHS Newborn Screening Program (NSP). HIV-1 antibody in newborn blood is maternal anti-HIV-1 IgG and indicates infection in the mother.[6] Following procedures for HIV-1 antibody detection in dried blood,[6,7] specimens were tested by enzyme immunoassay (EIA) and Western blot at the CHDS Viral and Rickettsial Disease Laboratory in Berkeley. Repeatedly EIA-reactive specimens with Western blot bands present at any two of p24, gp41, or gp120/160 were interpreted as positive.[8]

Analyses treated study subjects as a sample of childbearing women, representative of those recently giving birth; numbers of confirmed HIV-1 positive specimens were assumed to be distributed as binomial random variables. Multiple logistic regression analyses examined associations between HIV-1 antibody status and maternal race/ethnicity, region of residence, and age group.

Results

Nearly half of eligible births were delivered to San Francisco Bay Area and Los Angeles residents, with the remainder (50.8%) delivered to residents of "other California areas" (see Table 1). Although maternal age was missing for many NSP records, the survey distribution of known ages closely matched the maternal age distribution from 1988 California birth certificates.[9] For race/ethnicity, however, comparison with birth certificates suggested higher rates of missing data for both Black and White women in the survey.

Blood samples were tested for 135 808 births, or 99.2% of those eligible for the survey; specimens for 1053 births were missing or inadequate. Antibody to HIV-1 was confirmed in 101 specimens; 7 specimens with indeterminate Western blots (five p24 only, two p55 only) were interpreted as negative. Thus, there is an

Frank J. Capell, Duc J. Vugia, Virginia L. Mordaunt, William D. Marelich, Michael S. Ascher, George C. Cunningham, Stephen S. Arnon, and Kenneth W. Kizer are all with the California Department of Health Services; Alan I. Trachtenberg is with the University of California at San Francisco.

Requests for reprints should be sent to Frank J. Capell, Tobacco-Related Disease Research Program, Office of the President, University of California, 300 Lakeside Drive, Oakland, CA 94612-3550.

This paper was submitted to the journal November 8, 1990, and accepted with revisions June 10, 1991.

TABLE 1—HIV-1 Seroprevalence in California Childbearing Women: Third Quarter 1988

Maternal Characteristic	Number Tested	Number Positive	Rate (%)	95% CI	Odds Ratio	95% CI
Area of residence						
San Francisco Bay Area[a]	20 978	27	0.129	(.087–.190%)	2.96	(1.76–4.98)
Los Angeles	45 891	44	0.096	(.071–.130%)	2.20	(1.39–3.51)
Other California areas[b]	68 939	30	0.044	(.030–.063%)	1.00	
Race/ethnicity						
Asian	8588	3	0.035	(.009–.111%)	1.16	(0.34–4.00)
Black/Non-Hispanic	9369	34	0.363	(.255–.513%)	12.13	(6.89–21.99)
White/Non-Hispanic[b]	53 310	16	0.030	(.018–.050%)	1.00	
Hispanic	44 084	32	0.073	(.050–.104%)	2.42	(1.33–4.41)
Other/Unknown	20 457	16	0.078	(.046–.130%)	2.61	(1.30–5.21)
Age Group						
< 25 years	35 164	30	0.085	(.059–.124%)	2.32	(1.29–4.16)
25–34 years[b]	48 868	18	0.037	(.023–.059%)	1.00	
35+ years	8936	8	0.090	(.042–.184%)	2.43	(1.05–5.59)
Unknown	42 840	45	0.105	(.078–.142%)	2.85	(1.65–4.93)
Total sample	135 808	101	0.074	(.061–.091%)		

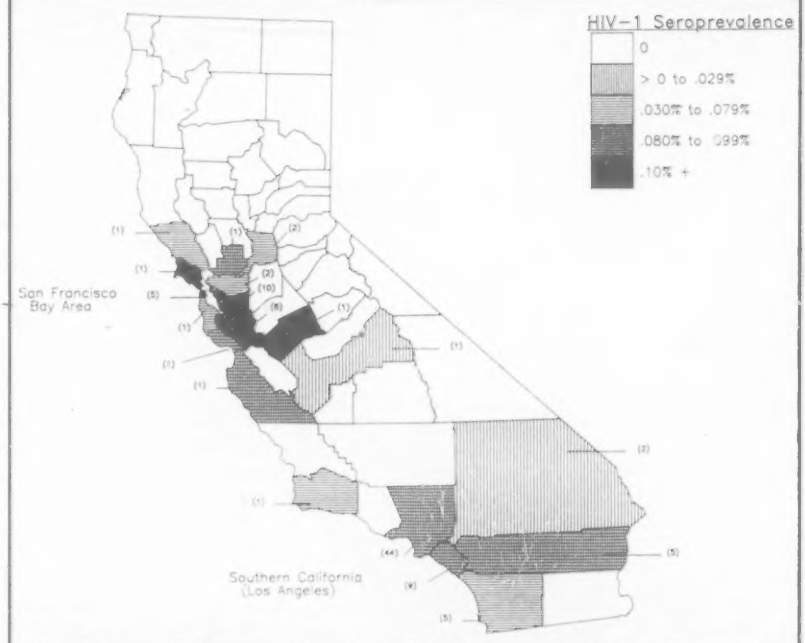
Note. CI = confidence interval.
[a]Includes Alameda, Contra Costa, Marin, San Francisco, San Mateo, and Santa Clara counties.
[b]Reference category for odds ratio calculation.

overall prevalence of infection of 0.074%, or 7.4 per 10 000 live births during the study period. Seroprevalence was lowest among women in the other California area (4.4 per 10 000). Los Angeles childbearing women were more than twice as likely to be infected as those from the other California area. San Francisco Bay Area women had the highest region-specific rate of HIV-1 infection, nearly three times that for other California women.

To provide additional geographic detail, Figure 1 displays rates and numbers of seropositive women for individual California counties. During the study period, one or more HIV-1 seropositive women gave birth in 19 of the state's 58 counties. Although the clustering of infection around San Francisco and Los Angeles suggests limited geographic spread, these 19 counties make up 86% of California's population.

Relatively low rates of infection were observed for Asian and White women in the survey (3.5 and 3.0 per 10 000, respectively). However, Hispanic women in the survey had just over twice the level of infection seen in White women, and Black women had the highest rate of infection: 36.3 per 10 000. In other words, 1 in 275 Black women delivering live babies during the study period was infected with HIV-1. This is over 12 times the rate for White women in the study. Seroprevalence among women with race/ethnicity coded as "other/unknown" was similar to that among the total sample.

Age-specific HIV-1 infection rates were lower among women under 25 years and those over 34 years in comparison with survey participants aged 25 to 34. These differences should be interpreted cautiously because age was more often missing among HIV-1-positive women. (Age was unknown in 45% of HIV-1-infected women vs 31.5% of women overall.)



Shaded areas indicate ranges of county-specific HIV-1 seroprevalence rates. Shown in parentheses is the number of HIV-1 seropositive childbearing women for each county with an infection rate greater than zero.

FIGURE 1—Geographic distribution of HIV-1 seroprevalence in California childbearing women, by county of residence: third quarter 1988.

In logistic regression analyses, both the elevated infection levels for Hispanic and Black women (relative to those for White and Asian women) and the differences among maternal age groups seen in Table 1 persisted after adjustment for possible confounding. With respect to re-

gional differences, only residence in the San Francisco Bay Area was independently associated with additional risk of infection. Interactive effects (e.g., being *both* Black *and* a San Francisco Bay Area resident) did not contribute to observed group differences in HIV-1 infection.

Discussion

As noted, HIV-1 infection was confirmed in 7.4 per 10 000 California childbearing women; this is in the same range as California female military applicants[10] and childbearing women in Texas,[5] but it is lower than findings on similar surveys in the eastern United States.[5,6,11] Preliminary results from a 1989 survey of California births indicate that, as in other states, overall seroprevalence has remained relatively stable. Within California, HIV-1 seroprevalence was highest among San Francisco Bay Area women; other regional variation appeared to reflect geographic differences in population demographic composition. AIDS cases had previously been reported in all counties where births to seropositive women occurred; one county, however, had no previous IDU or heterosexual contact cases in females, and three had no previous cases of "parent-at-risk" mode pediatric AIDS.[1]

More than a third of the estimated 400 HIV-1-infected California women delivering live babies annually are Black, and nearly another third are Hispanic. In contrast, whereas 32% of California female IDU and heterosexual contact AIDS cases are Black, only 15% are Hispanic. The proportion of Hispanic women among prevalent seropositives suggests that their representation among new female AIDS cases in the state is likely to increase, perhaps as much as twofold.[12]

This survey cannot address how sociodemographic differences in infection rates relate to patterns of risk behavior.[13] Although IDU and heterosexual contact each account for roughly equal numbers of female AIDS cases in California, the odds that a Black female AIDS case became infected through IDU (vs heterosexual contact) are three times greater than those for a Hispanic female case.[14] Such differences may reflect differing drug use or sexual mixing patterns, earlier introduction of HIV-1 into certain ethnic communities, or culturally linked differences in ascertaining transmission mode on AIDS case reports.

Over half a million children (532 708) were born in California during calendar year 1988;[9] assuming 25% to 35% vertical transmission efficiency,[2] the present survey suggests that roughly 120 of these infants were HIV-1 infected. This exceeds the cumulative number of "parent-at-risk" pediatric AIDS cases reported to date in California.[1]

Unlinked serosurveys in childbearing women may provide the best available information on the status of the HIV-1 epidemic among sexually active women in the United States. Increasingly, these surveys are being used to estimate HIV-1 infection rates or to model AIDS incidence in other population groups,[5,15] thus underscoring the need for careful assessments of survey bias.[16] The present survey covered 99.8% of live births during the study period, with HIV-1 serostatus determined for 99.2% of these. Demographic data were missing for many in the survey and, especially for maternal age, appeared less completely ascertained for women with higher rates of infection. Risk of HIV-1 infection may differ for eligible women excluded from the survey as well as for those who do not become pregnant or whose pregnancies do not end in live births. Additional data from population-based surveys of risk behavior and reproductive history and from HIV-1 serosurveys in other groups of women are needed to establish the limits of generalizability of information from surveys in childbearing women. □

Acknowledgments

This work was supported through Cooperative Agreement U62/CCU90219 between the California Department of Health Services (CDHS) and the Centers for Disease Control. The authors gratefully acknowledge the valuable contributions of Cynthia Cossen and Shirley Hagens of the CDHS Viral and Rickettsial Disease Laboratory; Daniel Morales and Terry Kennedy of the CDHS Clinical Chemistry Laboratory; and Erica Gordon, Karen Kan, and Wayne Nelson of the CDHS Genetic Disease Branch. Jim Singleton of the CDHS Office of AIDS provided statistical consultation. Erica Gordon, Richard Emmons, Jim Singleton, and Mary Jess Wilson provided useful comments on earlier drafts of the manuscript.

References

1. California Department of Health Services, Office of AIDS. AIDS surveillance report. *California AIDS Update.* 1990;3:24–25.
2. Oxtoby M. Perinatally acquired human immunodeficiency infection. *Pediatr Infect Dis J.* 1990;9:609–619.
3. Pappaioanou M, Dondero TJ, Peterson LR, Onorato IM, Sanchez CD, Curran JW. The family of HIV seroprevalence surveys: objectives, methods, and uses of sentinel surveillance for HIV in the United States. *Public Health Rep.* 1990;105:113–119.
4. Pappaioanou M, George JR, Hannon WH, et al. HIV seroprevalence surveys of childbearing women: objectives, methods, and uses of the data. *Public Health Rep.* 1990;105:147–152.
5. Gwinn M, Pappaioanou M, George JR, et al. Prevalence of HIV infection in childbearing women in the United States: surveillance using newborn blood samples. *JAMA.* 1991;265:1704–1708.
6. Hoff R, Berardi VP, Weiblen BJ, et al. Seroprevalence of human immunodeficiency virus among childbearing women: estimation by testing samples of blood from newborns. *N Engl J Med.* 1988;318:525–529.
7. George JR, Hannon WH, Jones WK, et al. Serologic assays for human immunodeficiency virus antibody in dried blood specimens collected on filter paper from neonates. Washington, DC: US Dept of Health and Human Services, Public Health Service; August 1989.
8. Centers for Disease Control. Interpretation and use of Western blot assay for serodiagnosis of human immunodeficiency virus type 1 infections. *MMWR.* 1989;38(suppl S–7):1–7.
9. California Department of Health Services. Birth profiles by county and ZIP code: California 1988. Data Matters 1990; Report 90–02003.
10. Centers for Disease Control. *Quarterly Summary of HIV Testing of US Military Recruit Applicants.* Atlanta, Ga: February 1990.
11. Novick LF, Berns D, Stricof R, et al. HIV seroprevalence in newborns in New York State. *JAMA.* 1989;261:1745–1750.
12. DeGruttola V, Fineberg HV. Estimating prevalence of HIV infection: considerations in design and analysis of a national seroprevalence survey. *J AIDS.* 1989;2:472–480.
13. Dawson JA. HIV in intravenous drug users. *N Engl J Med.* 1990;322:632. Letter.
14. Capell FJ, Vugia DJ, Mordaunt VL, Wilson MJ. Epidemiology of AIDS and HIV among California women. Sixth International Conference on AIDS; June 20–24, 1990; San Francisco, Calif. Abstract Th.C.703.
15. Pappaioanou M, Buehler J, Gwinn M. Relationship between cumulative incidence of AIDS and prevalence of HIV in women of reproductive age in the United States. Fifth International Conference on AIDS; June 5–8, 1989; Montreal, Canada. Abstract M.A.P.6.
16. Groves, RM. A total survey error approach to AIDS-related survey research. In: Fowler FJ, ed. *Conference Proceedings: Health Survey Research Methods.* Washington, DC: US Dept of Health and Human Services; 1989:265–270.

Cesarean Section Rates in Italy by Hospital Payment Mode: An Analysis Based on Birth Certificates

Roberto Bertollini, MD, MPH, Domenico DiLallo, MD, Teresa Spadea, PhD, and Carlo Perucci, MD

ABSTRACT

This study, based on birth certificate data from 1985 through 1987, investigated cesarean section (CS) rates in the Lazio region of Italy and their relationship with mode of hospital care payment. Use of abdominal delivery increased from 22.3% in 1985 to 24.3% in 1987. CS rates were highest (34.7%) in private hospitals. A marked variation in the use of CS was associated with mode of hospital care payment independently from other predictors of abdominal delivery. (*Am J Public Health.* 1992;82:257–261)

Introduction

Cesarean section (CS) deliveries have consistently increased in the past decades in Europe, North America, and the Pacific.[1] In the United States one out of four deliveries took place by CS in 1988,[2] and CS is the most common surgical procedure.[3] High CS rates are also observed in developing countries such as Brazil[4] and Puerto Rico.[5]

Several explanations and hypotheses have been offered to explain this impressive change in medical practice, such as a policy of subsequent cesarean delivery after a first cesarean, reluctance to deliver breech babies vaginally, and fear of malpractice suit.[6]

The use of CS for delivery is associated with nonclinical factors: socioeconomic status of women,[7] physician practice style,[8,9] physician work schedule,[10] and private or clinic care.[11] It has been shown in the United States that source of payment for obstetrical care is strongly associated with CS use, with women privately insured having the highest CS rate.[12] This observation was not confirmed in other countries with different health care systems.

In a previous preliminary report focusing on 1985 data regarding the Lazio region of Italy, we observed a high overall frequency of CS (22.5%) and a higher rate in private maternities (32.3%), where women pay all expenses for delivery.[13] To confirm this preliminary finding and investigate CS time trends, we analyzed birth certificate data for single infants born in 1985 through 1987 to women resident in the Lazio region. We compared the rates of CS by hospital payment mode after controlling, with logistic regression analysis, for the potential confounding effects of age, parity, birth weight, gestational age, presentation at birth, and day of delivery.

Materials and Methods

The study was based on an analysis of the birth certificates of single infants (live or stillborn) weighing more than 500 g and born to women resident in Lazio, a large region of central Italy, in the years 1985 through 1987. The largest city of the region is Rome, and the total regional population is about 5 100 000. Birth certificates are collected by the Medical Birth Registry coordinated by the Epidemiology Unit of the Regional Health Authority. Data on births are routinely linked to death certificates regarding deaths in the first year of life. The characteristics of the registry and the validation of data are described elsewhere.[14]

Hospital of birth was classified according to mode of medical care payment. There are three types of maternities in the Italian National Health Service (NHS): (1) NHS public hospitals, where medical care is provided free of any charge; (2) NHS private hospitals, where NHS pays a portion of the expenses (e.g., daily fees and medication) and the patient is responsible for physician fees and special services (e.g., individual room); and (3) private hospitals, where patients pay all expenses and no support is provided by NHS.

There are 96 maternities in the region (56 NHS public hospitals, 26 NHS private hospitals, and 14 private hospitals), with a total of about 47 000 births per year. All private hospitals are small units with a low number of births (e.g., the mean in 1986 was 191 births per year). Only 5 hospitals have more than 1500 births per year; 2 are teaching hospitals.

The initial cohort considered for the present study consisted of 139 284 single births; 2618 subjects were excluded (90 home births and 2528 with unknown information on mode of delivery and/or hospital of birth). The final cohort studied included 136 666 single hospital births. CS

The authors are with the Lazio Region Epidemiology Unit, Rome, Italy.

Requests for reprints should be sent to Roberto Bertollini, MD, MPH, Lazio Region Epidemiology Unit, Via di S. Costanza 53, 00198, Roma, Italy.

This paper was submitted to the journal April 30, 1991, and accepted with revisions September 16, 1991.

TABLE 1—Mode of Delivery by Year, Single Births, Lazio, Italy, 1985–1987

Mode of Delivery	1985		1986		1987	
	n	%	n	%	n	%
Vaginal	34 532	74.9	33 689	74.4	33 035	73.0
Cesarean section	10 276	22.3	10 312	22.8	11 025	24.3
Operative	1300	2.8	1278	2.8	1219	2.7

TABLE 2—Cesarean Section (CS) Rate by Hospital Payment Mode, Single Births, Lazio, Italy, 1985–1987

Hospital	Total Deliveries	CS Rate, %	Odds Ratio	95% Confidence Interval
NHS public	102 871	21.3	1.00	...
NHS private	25 692	26.7	1.34	1.30, 1.38
Private	8103	34.7	1.96	1.87, 2.06
Total	136 666	23.1	...	...

Note. See text for details on hospital classification. NHS = National Health Service.

rates were obtained by dividing the number of caesareans by the appropriate number of births (per 100).

Consistent with past research,[15–17] variables potentially predictive of CS were considered: maternal age, parity, gestational age, birth weight, presentation at birth, and day of the week of birth (weekday vs Sunday). The latter variable was considered because it had been associated with CS frequency in a previous analysis.[18] The Italian birth certificate does not include information on complications of pregnancy such as previous CS, dystocia, and fetal distress.

The strength of the association between CS rate, hospital payment, and the other predictive variables was evaluated using the odds ratio (OR); 95% confidence intervals (CIs) of the ORs were obtained using the test-based method.[19] Time trends were evaluated using the extended Mantel-Haenszel test.[20]

To assess the constancy of the relationship between payment source and CS rates, the data were stratified by maternal age, parity, birth weight, gestational age, presentation at birth, and day of delivery and the adjusted ORs calculated according to the Mantel-Haenszel method.[21] To further validate the findings derived from the above method, logistic regression analysis was employed to simultaneously estimate the independent effect of payment mode adjusted for the other predictive variables. The results of the multivariate analysis closely agreed with those of the tabular one. No statistical criteria (such as the stepwise procedure) were employed to select variables to be included in the model. Tests of significance for model selection depend on the size of the studied population, and, in large studies such as the present one, any trivial variable may have a highly significant association with the event under study (small P values) regardless of its epidemiological and clinical meaning.[22] Adjusted ORs and 95% CIs were derived from the estimated regression coefficient.[23] An interaction term (parity by maternal age) was included in the equation to take into account the joint effect of these two variables.

Results

The overall rate of CS in the years 1985 through 1987 among the 136 666 women delivering single infants in Lazio was 23.1% (Table 1). A marked statistically significant time trend ($P < .0001$) was observed over the time period studied (from 22.3% in 1985 to 24.3% in 1987). Provisional data for 1988 indicated a rate of 25.2% (R.B., unpublished data, 1991). Operative deliveries (forceps plus vacuum) constituted 2.8% of total deliveries without significant time trends.

A strong relationship was found between mode of hospital payment and CS rate. Women delivering in private hospitals had the highest CS rate (34.7%), while the rate of abdominal deliveries was lowest in NHS public hospitals (21.3%) (Table 2).

Maternal Age

The CS rate increased with maternal age: rates ranged from 16.4% among women younger than 20 years to 43.8% among those older than 39, an OR of 3.97 (95% CI = 3.54, 4.45). The rate was 18.2% among women 20 through 24, 21.7% among women 25 through 29, 26.4% among women 30 through 34, and 33.9% among women 35 through 39.

CS rates in private hospitals were significantly higher than those in NHS private and public hospitals in all age classes, except among the younger women (Figure 1). In the latter group, women delivering in NHS private institutions experienced the highest CS rate (18.3% vs 14.7% in NHS public and 16.6% in private hospitals); the difference among the three types of hospitals was of borderline statistical significance ($P = .055$). The CS rate among women older than 39 was extremely high in the private hospitals (51.1%).

Parity

The rate of CS significantly decreased with parity (23.9% for parity 0 and 22.4% for parity higher than 0). Within each parity class, the highest CS rate was observed in private hospitals (Figure 2). Since delivering in private hospitals can be judged as a proxy of the socioeconomic status of women,[18] and considering that most affluent women tend to be of lower parity and older maternal age,[7] the ORs of CS rates by hospital type were adjusted by maternal age and parity distribution. The adjusted OR_{mh} of CS in private hospitals was 1.58 (95% CI = 1.50, 1.66), as compared with 1.96 in the crude analysis; no important change was observed for NHS private hospitals ($OR_{mh} = 1.32$ vs $OR_{crude} = 1.34$).

Birth Weight and Gestational Age

The rates of CS were highest among women who delivered low-birth-weight (LBW) infants (34.3%), decreased with increasing birth weight (22.5%), and slightly increased for large babies (23.6%) (Table 3).

For each birth-weight class, CS rates significantly increased in going from NHS hospitals to the other two categories ($P < .0001$), although in private units a lower proportion of LBW and large babies at higher risk of CS were reported.

CS rates among women who delivered prematurely were highest (30.0%), decreased for full-term pregnancies, and increased again for postmature babies (24.5%) (Table 4). Again, rates were con-

sistently and significantly higher ($P < .0001$) in private units at any gestational age.

Complications

The use of CS is associated with various complications of pregnancy that are indications for carrying out the procedure.[24] The only complication reported in the Italian birth certificate is breech presentation of fetus at labor. The overall frequency of this complication was 3.3% and did not vary by hospital type. The rates of CS were highest in private and NHS private hospitals (84.7%), as compared with NHS public hospitals (75.7%) ($P < .001$).

Day of Delivery

CS rates were consistently lower on Sundays in all types of hospitals (overall CS rate of 15.7%), as compared with weekdays (overall rate of 24.1%). There was a statistically significant trend with regard to frequency of deliveries on Sundays ($P < .0001$): The lowest rate was observed in private hospitals (8.8%), the highest in NHS public units (12.2%), and an intermediate value in NHS private maternities (10%).

Multivariate Analysis

The results of the simultaneous multiple logistic adjustment are shown in Table 5. The ORs referring to maternal age and parity are shown for each combination of these two variables. Compared with the univariate analysis of unadjusted ORs, the multivariate assessment indicated that hospital payment mode was still associated with the likelihood of CS when allowance was made for the effects of the other variables present in the model. However, the OR associated with private hospitals was lower than the one observed in the crude analysis. This could be attributed to the maternal age–parity effect discussed previously, in part counterbalanced by the combined effect of the other variables in the model. The CS OR associated with NHS private maternities was practically identical to the one observed in the crude analysis.

The effect of maternal age on CS rate was marked for older/parity 0 women as compared with younger women and women of the same age at parity higher than 0.

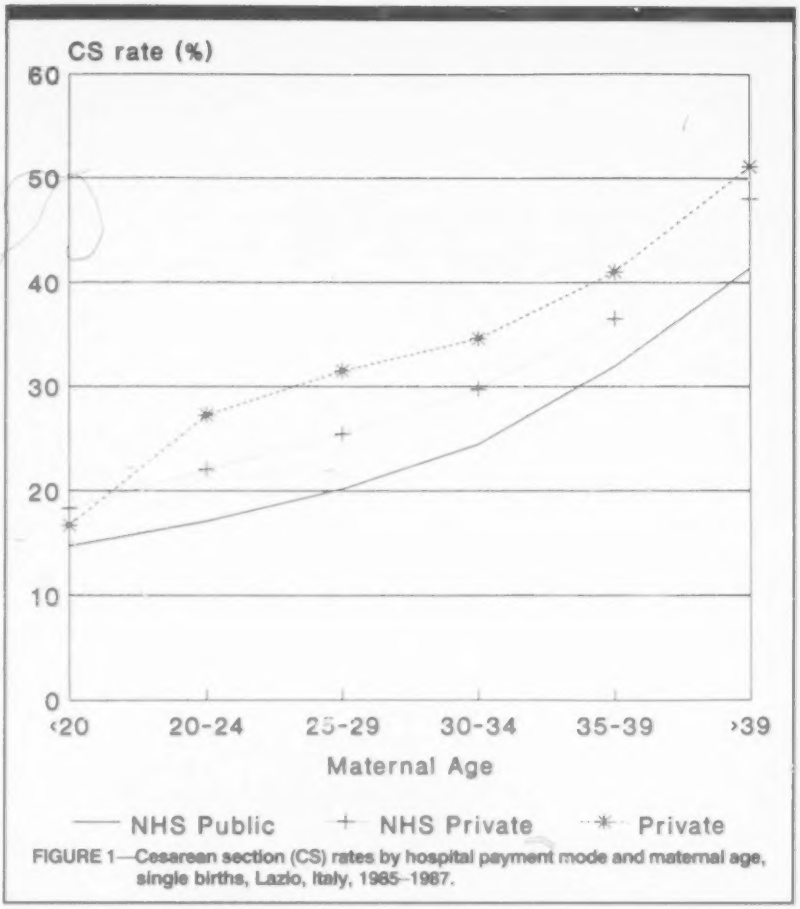


FIGURE 1—Cesarean section (CS) rates by hospital payment mode and maternal age, single births, Lazio, Italy, 1985–1987.

Discussion

This study showed that in the Lazio region of Italy, approximately one of four babies was born by CS in the time period 1985 through 1987, the highest rate reported in Europe so far. Lazio figures were comparable to those observed in the United States in the same time period,[25] and were largely beyond the figure considered acceptable by international agencies.[26] The proportion of operative deliveries (forceps and vacuum) was minimal and did not follow any time trend.

Mode of payment for obstetric care was strongly associated with CS rates after controlling for the effects of the other available determinants of abdominal delivery. Because of the limited information included in the Italian birth certificate, the effect of some known determinants of CS could not be taken into account, namely previous CS delivery, dystocia, and fetal distress. However, it is unlikely that these variables were differently distributed in the three types of hospitals to such an extent as to modify the observed effect of mode of payment on CS rates. In fact, the heterogeneous CS rates by payment mode were also observed in primiparous women (Figure 2). On the other hand, data referring to breech deliveries, the only complication of pregnancy reported in the birth certificate, confirmed higher CS rates in NHS private and private maternities than in NHS public hospitals.

The heterogeneity of CS rates by hospital payment mode might be attributable to the interaction of a variety of factors, including characteristics of patients, mix of physicians, and organizational features. In the Lazio region, delivering in private maternities can be judged as a proxy of high socioeconomic status,[18] which is associated with higher CS rates.[7] In addition, a multiplicity of financial incentives, such as a higher physician fee for CS as compared with vaginal delivery and congruence of physician interest with that of hospital (long hospital stay after CS), may influence directly or indirectly the clinical decision.[27] It is interesting to note that, in Italy, most obstetricians work part-time in both public hospitals and private institutions (NHS private or private), making the two health care settings fully comparable with regard to physician training and age. This suggests that different decision-making rules were applied to different health-

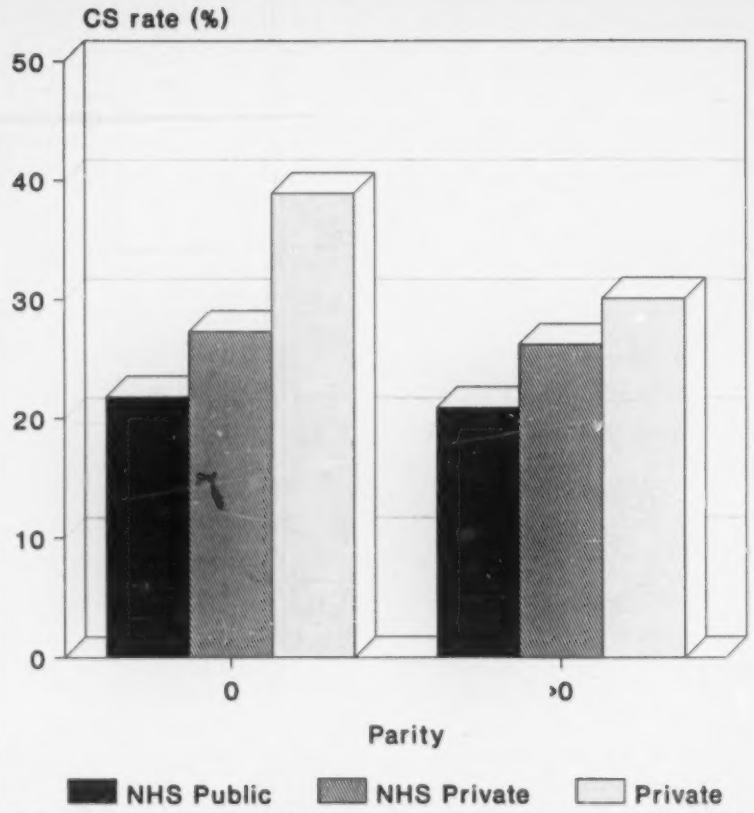


FIGURE 2—Cesarean section (CS) rates by hospital payment mode and parity, single births, Lazio, Italy, 1985–1987.

TABLE 3—Cesarean Section (CS) Rates by Birth Weight and Hospital Payment Mode, Single Births, Lazio, Italy, 1985–1987

	NHS Public		NHS Private		Private		Total
Birth Weight, g	No. of Deliveries	CS Rate, %	No. of Deliveries	CS Rate, %	No. of Deliveries	CS Rate, %	CS Rate, %
<2500	4911	33.6	915	37.5	283	47.0	34.3
2500–3999	89 720	20.6	22 712	26.2	7263	34.1	22.5
>3999	7961	21.9	1957	26.9	532	36.7	23.6

Note. NHS = National Health Service.

TABLE 4—Cesarean Section (CS) Rates by Gestational Age and Hospital Payment Mode, Single Births, Lazio, Italy, 1985–1987

	NHS Public		NHS Private		Private		Total
Gestational Age, wk	No. of Deliveries	CS Rate, %	No. of Deliveries	CS Rate, %	No. of Deliveries	CS Rate, %	CS Rate, %
<37	5643	29.3	1082	33.6	343	44.3	30.0
37–41	91 160	20.7	23 483	26.2	7430	34.2	22.6
>41	4256	23.4	867	30.0	232	34.9	24.5

Note. NHS = National Health Service.

TABLE 5—Odds Ratios and 95% Confidence Intervals of Cesarean Sections by Selected Determinants,[a] Single Births, Lazio, Italy, 1985–1987

Determinant	Odds Ratio	95% Confidence Interval
Hospital payment source		
NHS public	1.00	...
NHS private	1.35	1.31, 1.40
Private	1.64	1.56, 1.73
Birth weight, g		
<2500	1.47	1.38, 1.58
2500–3999	1.00	...
>3999	1.11	1.06, 1.16
Gestational age, wk		
Preterm (<37)	1.07	1.00, 1.14
At term (37–41)	1.00	...
Postterm (>41)	1.27	1.19, 1.36
Presentation		
Breech[b]	13.28	12.32, 14.32
Day of the week		
Weekday[c]	1.67	1.59, 1.75
Parity 0		
Maternal age, y		
<20	1.00	
20–24	1.23	1.11, 1.36
25–29	1.58	1.43, 1.74
30–34	2.29	2.07, 2.54
35–39	5.07	4.48, 5.73
>39	15.48	12.19, 19.66
Parity > 0		
Maternal age, y		
<20	0.84	0.62, 1.14
20–24	1.11	1.00, 1.24
25–29	1.33	1.21, 1.47
30–34	1.68	1.52, 1.85
35–39	2.17	1.95, 2.40
>39	2.83	2.46, 4.58

Note. NHS = National Health Service.
[a]Adjusted for all other factors present in the model.
[b]Cesarean section rate observed in all other presentations taken as reference.
[c]Sunday cesarean section rate taken as reference.

care settings for comparable clinical situations and comparable availability of diagnostic and surgical facilities. These different rules might be associated with the following factors: scheduling practices (such as 24-hour coverage by a staff obstetrician) that diminish the incentives for "convenience" CS in public maternities,[10] social congruity between obstetricians and middle-class parturients associated with the attitudes of these women toward the potential benefits of medical technology,[7] and the stronger fear of malpractice suits in cases of complication of pregnancy when caring for affluent and

more educated women, although malpractice suits in obstetrics are not common in Italy. All these elements may interact with physician practice style, which is associated with variability of CS use.[8] The decision to undertake a CS section is, in fact, not guided by unequivocal criteria; on the contrary, it is associated with a marked interobserver and, more surprising, intraobserver variability.[28]

In summary, our analysis of birth certificate data showed that the CS rate in 1985 through 1987 in the Lazio region was one of the highest ever reported. No tendency toward a leveling off of the rate, as observed in other countries,[2,5] was detected. This finding, together with the association of CS rate and hospital payment mode, was of great concern and underlines the urgent need for adopting regional and national strategies to reduce the use of abdominal delivery to more acceptable levels. □

Acknowledgments

We thank Dr. Pamela Pantano for her help in data collection and collation. We are also grateful to Dr. Francesco Forastiere and Dr. Marina Davoli for their useful comments on the first draft of the manuscript.

References

1. Notzon FC, Placek PJ, Taffel SM. Comparisons of national cesarean-section rates. *N Engl J Med.* 1987;316:386–389.
2. Taffel SM, Placek PJ, Moien M. 1988 U.S. cesarean-section rate at 24.7 per 100 births—a plateau? *N Engl J Med.* 1990;323:199–200. Letter.
3. Rutkow IM. Obstetric and gynecologic operations in the United States, 1979 to 1984. *Obstet Gynecol.* 1986;67:755–759.
4. Rodrigues J. Urban hospital cesarean section delivery rates in Paraiba State, Brazil, 1977–1981. *Am J Public Health.* 1988;78:704–705.
5. Notzon FC. International differences in the use of obstetric interventions. *JAMA.* 1990;263:3286–3291.
6. Taffel SM, Placek PJ, Liss T. Trends in the United States cesarean section rate and reasons for the 1980–85 rise. *Am J Public Health.* 1987;77:955–959.
7. Gould JB, Davey B, Stafford RS. Socioeconomic differences in rates of cesarean section. *N Engl J Med.* 1989;321:233–239.
8. Goyert GL, Bottoms SF, Treadwell MC, Nehra PC. The physician factor in cesarean birth rates. *N Engl J Med.* 1989;320:706–709.
9. De Mott RK, Sandmire HF. The Green Bay cesarean section study I. The physician factor as a determinant on cesarean birth rates. *Am J Obstet Gynecol.* 1990;162:1593–1599.
10. Fraser W, Usher RH, McLean FH, et al. Temporal variation in rates of cesarean section for dystocia: does 'convenience' play a role? *Am J Obstet Gynecol.* 1987;156:300–304.
11. Haynes de Regt R, Minkoff HL, Feldman J, Schwarz RH. Relation of private or clinic care to the cesarean birth rate. *N Engl J Med.* 1986;315:619–624.
12. Stafford RS. Cesarean section use and source of payment: an analysis of California hospital discharge abstracts. *Am J Public Health.* 1990;80:313–315.
13. Bertollini R, Di Lallo D, Rapiti E, Perucci CA. Cesarean section rates in Italy. *Am J Public Health.* 1987;77:1554. Letter.
14. Bertollini R, Di Lallo D, Perucci CA, et al. Description and validation of the natality and infant mortality information system in Lazio. *Neonatologica.* 1988;2:26–34.
15. National Institutes of Health. *Cesarean Childbirth: Report of the NICHD Task Force on Cesarean Childbirth.* Bethesda, Md: National Institutes of Health; 1981. DHHS publication no. (NIH) 82-2067.
16. Goldfarb MG. *Who Receives Cesareans: Patient and Hospital Characteristics.* Rockville, Md: National Center for Health Service Research; 1984. DHHS publication no. (PHS) 84-3345.
17. Williams RL, Chen PM. Controlling the rise in cesarean section rates by the dissemination of information from vital records. *Am J Public Health.* 1983;73:863–867.
18. Bertollini R, Di Lallo D, Di Lena P, Gittarelli D, Papini P, Perucci CA. Rapporto del Sistema di Sorveglianza della natalitáe mortalitá infantile nel Lazio. Anno 1985–86. *Progetto Salute.* 1990;14:1–178.
19. Miettinen O. Estimability and estimation in case-referent studies. *Am J Epidemiol.* 1976;103:226–235.
20. Schlesselman JJ. *Case Control Studies. Design, Conduct, Analysis.* New York, NY: Oxford University Press; 1982:203–206.
21. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst.* 1959;22:719–748.
22. Schlesselman JJ. *Case Control Studies. Design, Conduct, Analysis.* New York, NY: Oxford University Press; 1982:252–254.
23. Woolson RF. *Statistical Methods for Biomedical Research.* New York, NY: John Wiley and Sons; 1987.
24. Mann LI, Gallant J. Modern indications for cesarean section. *Am J Obstet Gynecol.* 1979;135:437–441.
25. Myers SA, Gleicher N. 1988 U.S. cesarean-section rate: good news or bad? *New Engl J Med.* 1990;3233:200. Letter.
26. World Health Organization. Appropriate technology for birth. *Lancet.* 1985;ii:436–437.
27. Stafford RS. The impact of nonclinical factors on repeat cesarean section. *JAMA.* 1991;265:59–63.
28. Barrett JFR, Jarvis GJ, MacDonald HD, Buchan PC, Tyrrell SN, Lilford RJ. Inconsistencies in clinical decisions in obstetrics. *Lancet.* 1990;336:549–551.

Professional Relationships between Midwives and Physicians: Collaboration or Conflict?

Laura-Mae Baldwin, MD, MPH, Heidi L. Hutchinson, MD, and Roger A. Rosenblatt, MD, MPH

ABSTRACT

This study examines the professional relationships between midwives and physicians providing obstetrical care in Washington State. Four hundred ninety-six randomly sampled family physicians and obstetrician-gynecologists and 211 certified nurse, licensed, and lay midwives were surveyed to learn more about midwife/physician consulting relationships. Only certified nurse midwives have forged mutually satisfactory relationships with the physician community. Increased hospital-based training and practice opportunities are needed before licensed midwives can improve their professional relationships with physicians. (*Am J Public Health.* 1992;82:262–264)

Introduction

Midwives have become increasingly important birth attendants in Washington State, particularly as many family physicians and obstetricians have discontinued their obstetrical practices because of malpractice liability concerns.[1] The proportion of births attended by licensed and certified nurse midwives increased more than fortyfold between 1975 and 1989, from less than 1% to 4.6% of all births (P. Starzyk, personal communication, May 1991).[2] In 1989, 3436 Washington birth records listed a midwife as the attendant. Since 1988, six clinics, each designed to serve over 300 low-income pregnant women per year, have been established in Washington using certified nurse midwives as providers. One barrier to the further growth of midwifery has been the development of functional, mutually satisfying professional relationships between midwives and physicians. We surveyed Washington's obstetricians, family physicians, and midwives to better understand the nature of their professional relationships.

Methods

We identified 64 certified nurse midwives and 68 licensed midwives registered with Washington's Department of Licensing in 1987, as well as 79 lay midwives—66 identified by a Christian lay midwifery organization and 13 by the Department of Licensing. Because lay midwives are not licensed, our sample may not be truly representative of this group. Certified nurse midwives are registered nurses with advanced training in midwifery in a program accredited by the American College of Nurse-Midwives.[3] Licensed midwives consist mainly of graduates of the Seattle Midwifery School, a 27-month program for students who have completed at least 2 years of college.[4] Washington is unusual in having a large group of professional, nonnurse midwives who are recognized and licensed by the state. Lay midwives are allowed to practice in Washington State without any specific educational requirements but cannot advertise or charge for their services.[2] We also identified 496 family physicians and obstetrician-gynecologists randomly sampled from the 1988 American Medical Association directory's listing of Washington physicians.[5]

To describe the geographic distribution of midwives and physicians, we aggregated Washington's 39 counties into 4 categories. Urban counties are those with 76% to 100% of the population living in an urban area, semiurban 51% to 75%, semirural 25% to 50%, and rural 0% to 25%. In Washington State, 4 counties are classified as urban, 15 as semiurban, 8 as semirural, and 12 as rural.

Two survey mailings were sent to midwives and physicians in the summer of 1988 asking questions regarding physician/midwife demographics, practice characteristics, and consultation and referral relationships. After excluding those providers who had retired, died, or recorded a different specialty, 51% of physicians (234/458), 66% of certified nurse midwives (42/64), 71% of licensed midwives (48/68), and 42% of lay midwives (33/78) responded to the questionnaire.

Results

Ninety percent of obstetricians (45/50), 37% of family physicians (66/180), 69% of certified nurse midwives (29/42), 71% of licensed midwives (34/48), and 48% of lay midwives (16/33) were providing some component of obstetrical care in 1988. Nearly all practicing midwives provided prenatal care, but certified nurse

Laura-Mae Baldwin and Roger A. Rosenblatt are with the Department of Family Medicine, University of Washington, Seattle. Heidi L. Hutchinson is with the Swedish Hospital, Seattle, Washington.

Requests for reprints should be sent to Laura-Mae Baldwin, MD, MPH, Department of Family Medicine, University of Washington, RF-30, Seattle, WA 98195.

This paper was submitted to the journal June 26, 1990, and accepted with revisions July 18, 1991.

TABLE 1—Practice Characteristics of Obstetrical Care Providers

	Physicians		Midwives		
	Obstetricians (n = 45)	Family physicians (n = 66)	Certified nurse (n = 29)	Licensed (n = 34)	Lay (n = 16)
Average no. of deliveries in 1987	142	39	83	38	15
Average years in practice	13	12	8	7	7
Private practice, %	80	79	31	97	99

TABLE 2—Practice Location of Obstetrical Care Providers

	Physicians		Midwives		
Practice location, %	Obstetricians (n = 45)	Family physicians (n = 66)	Certified nurse (n = 29)	Licensed (n = 34)	Lay (n = 16)
Urban	62	51	69	54	7
Semiurban	27	29	19	27	65
Semirural	11	11	4	19	14
Rural	0	9	8	0	14

midwives (86%) and lay midwives (81%) attended deliveries to a lesser extent than licensed midwives (97%). Obstetricians and certified nurse midwives attended the greatest number of deliveries (Table 1). Most physicians and midwives practiced in urban and semiurban counties (Table 2), while few lay midwives practiced in urban counties.

Certified nurse midwives (69%) worked primarily as employees of hospitals or physicians, while licensed and lay midwives were almost exclusively in private practice. Certified nurse midwives attended the majority of their deliveries in birthing centers or hospitals (76%), while licensed and lay midwives attended deliveries almost exclusively in the home (91% and 99%, respectively).

Nearly 67% of obstetricians and 44% of family physicians had consulted for or received a referral from a midwife at some point in their medical practice. Fifty-six percent of obstetricians and 32% of family physicians practicing obstetrics had an existing referral relationship with a midwife in 1987. Of those who had not worked with a midwife, 50% of obstetricians and 51% of family physicians stated that they would be willing to do so. Seventy-six percent of all physicians practicing obstetrics felt that midwives can provide an acceptable alternative to physician care for low-risk pregnancies. No physician was willing to attend a home delivery, however.

Of the practicing midwives, 100% of certified nurse midwives (29), 97% of licensed midwives (33), and 50% of lay midwives (8) had an existing referral relationship with a physician. Certified nurse midwives predominantly consulted with and referred to obstetricians, while licensed and lay midwives had referral relationships with both obstetricians and family physicians (Table 3). Referral relationships for certified nurse midwives generally were established by their place of employment or the consulting physician (83%), while licensed and lay midwives usually approached physicians directly. Sixty-two percent of certified nurse midwives had a written contract; 52% met regularly with a physician to maintain their professional relationships.

TABLE 3—Midwife/Physician Relationships

	Certified Nurse Midwives (n = 29)	Licensed Midwives (n = 33)	Lay Midwives (n = 8)
Referral relationship with, %:			
Obstetrician	89	66	38
Family physician	11	34	38
Naturopath	0	0	25
Midwife-initiated relationship, %	17	88	88
Length of relationship, y	5	4	3
Written contract with physician, %	62	6	0
Regularly scheduled meetings/contact with physician, %	52	0	13
Average miles to physician practice	5	8	23

Sixty-nine percent of certified nurse midwives felt that no changes were needed in their referral relationships. Virtually all licensed midwives, however, wanted to establish more formal referral relationships with physicians. Midwives sought increased respect, appreciation, and communication from their physician colleagues. They also wanted greater integration with the traditional medical system through support for hospital privileges, more physicians willing to provide backup, and insurance coverage for home deliveries.

Discussion

Midwives are not a homogeneous group. Certified nurse midwives are closely aligned with traditional obstetrical services. They perform large volumes of deliveries, usually within hospitals in urban and semiurban counties. Lay midwives are most distant from the established medical community. Those we were able to contact either had no professional relationship or had informal, short-term relationships with physicians. Licensed midwives occupy an uncomfortable middle ground. They deliver almost exclusively in the home, yet rely on physicians—who, as a rule, refuse

to perform or assist in home deliveries—for consultation and referral. This group has developed less formal and less satisfying relationships with physicians than have certified nurse midwives.

Certified nurse midwives have developed a role in the institutionally based obstetrical care system, through which they are helping address problems of access for economically disadvantaged populations in Washington. Licensed and lay midwives have practiced primarily outside the mainstream of medicine. Without formal training and licensing, lay midwives will continue to practice in this manner. Licensed midwives, despite their status as licensed birth attendants, have been dissatisfied with their consulting relationships with physicians. Hospital-based training and practice opportunities must be developed if licensed midwives are to play a more significant role in providing obstetric care in Washington State. □

References

1. Rosenblatt RA, Detering B. Changing patterns of obstetric practice in Washington State: the impact of tort reform. *Fam Med.* 1988;20:101–107.
2. Starzyk PM. *Alternative Maternity Services in Washington State: Trends and Current Practice.* Olympia, Wash: Department of Social and Health Services; Health Data Unit; 1987.
3. Heaney CT, Malter A. *Midwifery Outside of the Nursing Profession: The Current Debate in Washington.* Seattle, Wash: University of Washington School of Public Health and Community Medicine, Health Policy Analysis Program; 1980.
4. Seattle Midwifery School. *Seattle Midwifery School Catalog, 1990–92.* Seattle, Wash: Seattle Midwifery School; 1989;9.
5. American Medical Association. *American Medical Directory: Physicians in the United States, 30th Edition.* Chicago, Ill: American Medical Association; 1986.
6. US Bureau of the Census. *County and City Data Book.* Washington, DC: US Government Printing Office; 1983:xxv.

Factors Associated with Artificial Feeding in Shanghai

Yue Chen, MD, PhD

ABSTRACT

Factors associated with artificial feeding were analyzed for 3285 infants in Shanghai. Boys, those from more highly educated families, and those born by assisted delivery or by cesarean section were more likely to be artificially fed than were girls, those from less educated families, and those born by spontaneous delivery. Infants whose birth weight was around 3750 g had the lowest probability of artificial feeding; higher and lower birth weights were positively associated with artificial feeding. (*Am J Public Health.* 1992;82:264–266)

Introduction

Artificial feeding has been documented as an important risk factor in increased mortality and morbidity from respiratory illness, gastroenteritis, and otitis.[1–4] In China, pneumonia and diarrhea are the most common diseases in infants and children.[5] Because breast-feeding makes a unique, fundamental contribution to the health of infants, the reasons for unsuccessful breast-feeding are important concerns.[6,7]

Factors found to be related to artificial feeding have been diverse and inconsistent. The amount of information available for some geographic areas is sparse; in particular, very little information is available for China, the world's most populous country.

In Shanghai, one third of the mothers studied did not initiate breast-feeding.[3,4] The proportion of artificial feeding was surprisingly high compared with the findings from other developing countries or even some developed countries.[6–8] This paper presents information on some factors associated with the infant-feeding practices of mothers in Shanghai.

Methods

This analysis was performed on the combined data from the Jing-An and Chang-Ning epidemiological studies of children's health in Shanghai.[9,10] A detailed description of the populations and methodology has been published elsewhere.[9,10] All together, the data of 3285 infants from these two studies were used in this analysis.

The relationship between birth weight and the proportion of artificial feeding was U-shaped. Nonlinear regression[11] was used to estimate parameter values for the model

$$\text{artificial feeding} = a + b\,(\text{birth weight} - c)^2.$$

A multiple logistic regression model was used to examine the individual contribution of each risk factor to proportion of artificial feeding. The transformed values of birth weight, $(\text{birth weight} - c)^2$, were used to fit the model. All other variables were treated categorically.

The author is with the Department of Medicine, University Hospital, Saskatoon, Saskatchewan, Canada.

Requests for reprints should be sent to Yue Chen, MD, PhD, Department of Medicine, Wing 3E, University Hospital, Saskatoon, Saskatchewan, Canada S7N 0X0.

This paper was submitted to the journal August 27, 1990, and accepted with revisions June 19, 1991.

Results

Table 1 shows that the proportion of artificial feeding was slightly higher in the Jing-An study than in Chang-Ning study. In Jing-An District, a higher proportion of parents had university education, the mothers were older at the time of birth, the infants were more likely to be delivered by cesarean section, and the average income per person was higher than that in Chang-Ning District. There were no significant differences between the two infant populations in sex, multiple birth, maternal history of spontaneous abortion, or average birth weight, which was 3229 g in both groups.

Of the 3285 infants, 1132 (34.5%) had never been breast-fed. As Table 2 shows, the proportion of artificial feeding was higher in boys than in girls. Artificially fed infants were likely to be from families in which the father was more highly educated than the general population. Infants who were born by assisted delivery or by cesarean section were more likely to be artificially fed than were those delivered by spontaneous labor. Multiple births had a higher proportion of artificial feeding than did the general population.

Birth weight was another important predictor of frequency of artificial feeding. The average birth weight for all infants was 3229 g (SD 441 g). The relationship estimated by nonlinear regression is shown in Figure 1 ($t = 6.453$, $P < .001$). Infants whose birth weight was 3750 g had the lowest proportion of artificial feeding (30.4%). Higher and lower birth weights were associated with higher proportions of artificial feeding.

The variables in the final logistic model included study (Jing An or Chang-Ning), sex, birth weight, type of delivery, and father's education. The adjusted odds ratios for the proportion of artificial feeding were 0.85 (95% confidence interval [CI], 0.73–0.98) for girls compared with boys and 0.76 (95% CI, 0.61–0.96) for those from families with less educated fathers compared with those from families with university-educated fathers. The adjusted odds ratios for children born by cesarean section and those born by assisted delivery were 1.45 (95% CI, 1.19–1.75) and 1.41 (95% CI, 1.14–1.74), respectively, compared with those born by spontaneous delivery.

The transformation of birth weight, $(\text{birth weight} - 3750)^2$, was used to fit the logistic model. The U-shaped relationship was slightly stronger after adjustment for the covariates.

TABLE 1—Comparison of Data Obtained from the Jing-An Study (n = 1058) and the Chang-Ning Study (n = 2227)

	Number (%) from Jing-An	Number (%) from Chang-Ning	Total
Sex			
Male	551 (52.1)	1155 (51.9)	1706 (51.9)
Female	507 (47.9)	1072 (48.1)	1579 (48.1)
Multiple birth			
Yes	25 (2.4)	43 (1.9)	68 (2.1)
No	1033 (97.6)	2184 (98.1)	3217 (97.9)
Feeding			
Breast or mixed	670 (63.3)	1483 (66.6)*	2153 (65.5)
Artificial	388 (36.7)	744 (33.4)	1132 (34.5)
Father's education			
University	154 (14.6)	200 (9.0)**	354 (10.8)
Secondary	883 (83.5)	1999 (89.8)	2882 (87.7)
Primary	21 (2.0)	28 (1.3)	49 (1.5)
Maternal age at birth			
< 30 y	768 (72.6)	1714 (77.0)**	2482 (75.6)
≥ 30 y	290 (27.4)	513 (23.0)	803 (24.4)
Mother's education			
University	73 (6.9)	76 (3.4)**	149 (4.5)
Secondary	966 (91.3)	2128 (95.6)	3094 (94.2)
Primary	19 (1.8)	23 (1.0)	42 (1.3)
Type of delivery			
Spontaneous	762 (72.0)	1501 (67.4)**	2263 (68.9)
Assisted	150 (14.2)	298 (13.4)	448 (13.6)
Cesarean section	146 (13.8)	428 (19.2)	574 (17.5)
History of spontaneous abortion			
Yes	97 (9.2)	206 (9.3)	303 (9.2)
No	961 (90.8)	2021 (90.7)	2982 (90.8)
Average income per person			
< 25 yuan	37 (3.5)	23 (1.0)**	60 (1.8)
25–49 yuan	967 (91.4)	1932 (86.8)	2899 (88.2)
≥ 50 yuan	54 (5.1)	272 (12.2)	326 (9.9)

*$.05 < P < .10$.
**$P < .01$.

The average birth weight in the multiple-birth group was significantly lower than that in the general population (2625 g vs 3242 g, $t = 11.654$, $P < .001$). After adjustment for the covariates, the effect of multiple birth is no longer statistically significant. No significant relationships were found between the proportion of artificial feeding and other factors, including maternal age at birth, mother's educational status, maternal history of spontaneous abortion, and average income per person in the family.

Discussion

The lowest proportion of artificially fed infants was found among those with birth weights of around 3750 kg. Both smaller and larger infants were more likely to be fed artificially. Previous studies have reported that lower birth weight was related to a lower proportion and shorter duration of breast feeding,[7,8] but the results from different studies are not consistent.

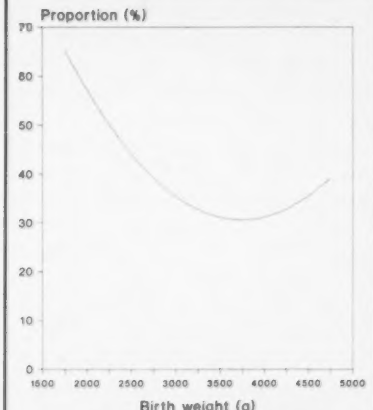


FIGURE 1—Relationship between proportion of artificial feeding (PAF) and birth weight (BWT).

The U-shaped relationship between birth weight and proportion of artificial feeding has not been described previously, possibly because in previous stud-

TABLE 2—Relationships between Risk Factors and Proportion of Artificial Feeding (AF)

	Number	AF (%)	χ^2 (*P* value)
Study area			
Jing-An	1058	36.7	3.385
Chang-Ning	2227	33.4	(0.066)
Sex			
Male	1706	35.9	3.407
Female	1579	32.9	(0.065)
Multiple birth			
Yes	68	52.9	10.501
No	3217	34.1	(0.001)
Father's education			
University	354	40.4	6.190
Secondary or primary	2931	33.7	(0.013)
Maternal age at birth			
< 30 y	2482	34.3	0.079
≥ 30 y	803	34.9	(0.779)
Mother's education			
University	149	32.9	0.171
Secondary or primary	3136	34.5	(0.679)
Type of delivery			
Spontaneous	2263	32.1	18.218
Assisted	448	39.7	(< 0.001)
Cesarean section	574	39.7	
History of spontaneous abortion			
Yes	303	37.6	1.480
No	2982	34.1	(0.224)
Average income per person			
< 25 yuan	60	35.0	1.050
25–49 yuan	2899	34.7	(0.591)
≥ 50 yuan	326	31.9	

ies, birth weight was treated dichotomously or categorically, and infants with birth weights of 3500 g or more were all classified into same group. The biological explanation for the U-shaped relationship remains unclear. It could be that small infants have difficulty in establishing breast feeding and large infants are unable to have their need for more milk satisfied by their mothers.

Artificial feeding was more common in infants born by assisted delivery or by cesarean section than in those born by spontaneous delivery. These results are consistent with the findings reported previously.[7,8,12,13] It has also been demonstrated that assisted delivery and cesarean section were associated with delay in starting breast-feeding.[14] Poor condition of abnormally delivered infants and the separation of mother and infant have been suggested as factors responsible for the delayed initiation of breast-feeding.[12] Delayed lactation and the use of anesthetics could have long-term consequences on the establishment and duration of breast-feeding.

In this analysis, it was found that paternal education was positively associated with the proportion of artificial feeding. It seems that in Shanghai, infants in well-educated families were more likely to be fed artificially. Some previous studies have found that lower maternal educational level was associated with a higher proportion of breast-feeding.[15,16] Other studies, however, showed that higher maternal education could improve breast-feeding practice.[7,8] The association seems to be linked to a variation by country or by ethnic group within a country.

In conclusion, birth weight, type of delivery, paternal education, and sex of the infant are significant predictors of artificial feeding in Shanghai. These findings will assist the development of breast-feeding promotion programs in Shanghai. □

Acknowledgments

The author thanks Dr Linda Pederson for her comments and Krista Baerg for her assistance in manuscript preparation.

References

1. Kovar MG, Serdula MK, Marks JS, Fraser DW. Review of the epidemiologic evidence for an association between infant feeding and infant health. *Pediatrics.* 1984;74:615–638.
2. Jason JM, Nieburg P, Marks JS. Mortality and infectious disease associated with infant-feeding practices in developing countries. *Pediatrics.* 1984;74:702–727.
3. Chen Y, Yu SZ, Li WX. Artificial feeding and hospitalization in the first 18 months of life. *Pediatrics.* 1988;81:58–62.
4. Chen Y. Synergistic effect of passive smoking and artificial feeding on hospitalization for respiratory illness in early childhood. *Chest.* 1989;95:1004–1007.
5. Zhou HK. Medical progress in China. *Chin Med J.* 1989;102:575–578.
6. Notzon F. Trends in infant feeding in developing countries. *Pediatrics.* 1984;74:648–666.
7. Simopopoulos AP, Grave GD. Factors associated with the choice and duration of infant-feeding practice. *Pediatrics.* 1984;74:603–614.
8. Forman MR. Review of research on the factors associated with choice and duration of infant feeding in less developed countries. *Pediatrics.* 1984;74:667–694.
9. Chen Y, Li WX, Yu SZ. Influence of passive smoking on admissions for respiratory illness in early childhood. *Br Med J.* 1986;293:303–306.
10. Chen Y, Li WX, Yu SZ, Qian WH. Chang-Ning Epidemiological Study of Children's Health: I. Passive smoking and children's respiratory diseases. *Int J Epidemiol.* 1988;17:348–355.
11. *SPSS-X™ User's Guide.* Chicago, Ill: SPSS Inc; 1988.
12. Tamminen T, Verronen P, Saarikoski S, Göransson A, Tuomiranta H. The influence of perinatal factors on breast feeding. *Acta Paediatr Scand.* 1983;72:9–12.
13. Palmer SR, Avery A. The influence of obstetric procedures and social and cultural factors on breast-feeding rates at discharge from hospital. *J Epidemiol Community Health.* 1979;33:248–252.
14. Martin J. Infant feeding 1975: attitudes and practice in England and Wales. HMSO: London; 1978.
15. Butz WP, DaVanzo J. *Determinants of Breastfeeding and Weaning Patterns in Malaysia.* Santa Monica, Calif: RAND Corp; 1981.
16. Zurayk HC, Shedid HE. The trend away from breast feeding in a developing country: a women's perspective. *J Trop Pediatr.* 1981;27:237–244.

Smoking-Related Behavior, Beliefs, and Social Environment of Young Black Women in Subsidized Public Housing in Chicago

Clara Manfredi, PhD, Loretta Lacey, RN, DrPH, Richard Warnecke, PhD, and Marianne Buis, PhD

ABSTRACT

Survey data indicate that young Black female smokers living in public housing are heavier smokers and have weaker motivation to quit, health beliefs and social environment less conducive to cessation, and less knowledge of where to get help to quit than other young Black female smokers in metropolitan Chicago. Compared with White women, the latter, other Black women smoke fewer cigarettes daily and have a stronger desire to quit and more concern about health reasons for quitting, but have a weaker belief in the risk of lung cancer from smoking, greater concern about quitting difficulties, and less knowledge of where to get help to quit. Low education, not race, is associated with higher smoking prevalence and less social pressure to quit or support for quitting. (*Am J Public Health.* 1992;82: 267–272)

Introduction

Smoking prevalence rates among young Black women of low socioeconomic status are high[1–3] and projected to remain so on the basis of current smoking cessation trends.[4–6] Smoking cessation programs tailored to this population are needed,[3,7] but few data are available about smoking differences within the Black population that are attributable to socioeconomic factors.

Smoking cessation may be particularly difficult to achieve among individuals with several of the characteristics associated with high smoking rates (including low education, income below poverty level, unemployment, employment in service-level jobs, and unmarried status[5,7,8]) and among those living in urban communities who have been described as overwhelmed by problems of poverty, crime, and decaying environment.[9–12] Under these cumulative conditions, beliefs and social norms favoring smoking cessation are likely to be weak and not very salient.

In this paper we explore the smoking-related beliefs and social environment of such a population and compare them with those of a general population of smokers. Survey data are presented for female smokers aged 18 to 39 in three groups: (1) Black residents of Chicago subsidized public housing ("Black SPH Residents"), a population subgroup of known low socioeconomic status[13]; (2) Blacks from the general metropolitan Chicago population ("Other Blacks") who are not SPH residents; (3) Whites from the general metropolitan Chicago population ("Whites") who are not SPH residents.

Methods

Samples

As part of the evaluation of a televised smoking cessation intervention, baseline telephone interviews were conducted with a general population sample of 2398 smokers who watched the evening news at least three times per week. This was a random-digit-dialing sample of the noninstitutionalized population, 18 years of age or older, in the Chicago metropolitan area.[14] The interviews were conducted 6 months before the intervention. The sample yielded 117 Other Black females and 496 White females aged 18 to 39. The SPH sample was identified by screening a random sample of 1849 apartments in three public housing developments. Each development and its buildings were sampled with probability proportionate to its share of the three developments' total population. Of these, 19% were unoccupied, the occupants of 9% refused or were unavailable, and 58% had no smokers qualifying by sex, age, or TV viewing habits.* Face-to-face interviews were completed with 246 (96%) of the 256 eligible respondents.

Table 1 shows education level, income, and employment, marital, and single-parenthood status for the three study groups.

Hypotheses

Hypothesis #1: Because of the high smoking prevalence rates associated with low socioeconomic status, Black SPH Residents are expected to be heavier

*About 87% of the general population sample and 89% of the SPH sample who were contacted and who otherwise qualified for the study were regular viewers of the local evening news, either between 4 PM and 6:30 PM or at 10 PM.

Clara Manfredi and Loretta Lacey are with the Prevention Research Center and Richard Warnecke is with the Survey Research Laboratory, all at the University of Illinois at Chicago. Marianne Buis is with the Department of Health Behavior, University of Alabama at Birmingham.

Requests for reprints should be sent to Clara Manfredi, PhD, Prevention Research Center, University of Illinois at Chicago, 850 W Jackson Blvd, Suite 400, Chicago, IL 60607.

This paper was submitted to the journal October 23, 1990, and accepted with revisions July 18, 1991.

TABLE 1—Demographic Characteristics, by Three Study Groups

	Black SPH Residents (n = 246) %	Other Blacks (n = 117) %	Whites (n = 496) %	% Differences (95% CI) Black SPH Residents vs Other Blacks	% Differences (95% CI) Other Blacks vs Whites
Education					
Less than high school	42	19	7	24 (14, 34)	12 (4, 20)
Completed high school	33	26	40	7 (−4, 17)	−14 (−24, −5)
More than high school	25	56	53	−30 (−41, −19)	3 (−7, 13)
Income					
Less than $13,000	100[a]	29	9	71 (60, 83)	20 (11, 29)
$13,000 to $25,000	0	37	22	−37 (−50, −25)	15 (2, 28)
More than $25,000	0	34	69	−34 (−46, −22)	−35 (−48, 22)
Employed full- or part-time	10	68	72	−58 (−67, −48)	−4 (−13, −5)
Married or living with partner	24	38	64	−14 (−24, −4)	−26 (−36, −16)
Single parent[b]	66	30	10	36 (25, 47)	20 (11, 29)

Note. SPH = subsidized public housing; CI = confidence interval.
[a]Not asked in interview. A median yearly income of $5,000 was reported in the Chicago Housing Authority's *1983 Statistical Report.*
[b]A single parent has at least one child in the household and is currently neither married nor living with a partner.

smokers and to have weaker motivation to quit smoking, fewer beliefs consistent with smoking cessation, weaker social pressure, and less social support to quit than Other Blacks.

Hypothesis #2: Because of the increasing smoking rates among young Black females in general (and regardless of the well-known smaller daily cigarette consumption among Blacks[15]), Other Blacks are expected to have weaker motivation to quit smoking, fewer beliefs consistent with smoking cessation, weaker social pressure, and less social support to quit than Whites.

Variables and Statistical Methods

We used 45 dependent variables, listed in Tables 2 and 3, to test the hypotheses above and to reflect psychosocial factors commonly appearing in the smoking cessation literature.[16–18] For each variable, differences between Black SPH Residents and Other Blacks and between Other Blacks and Whites are presented with 95% confidence intervals.

Each dependent variable in Tables 4 and 5 was regressed in simultaneous logistic regression[19] with education (a three-level category) and study group. This was done separately for the two group comparisons because (1) the patterns of differences by study groups, shown in tables 2 and 3, do not warrant treating the three groups as categories in an ordinal variable and (2) for many variables the effect of education is different or even reversed by study group. Adjusted and unadjusted odds ratios and 95% confidence intervals[20] were calculated for the group comparisons and for higher versus lower education. Results are presented only for variables for which the group comparison is significantly ($\alpha = .05$) altered by education or for which education itself has a significant effect.

Results

Regarding Hypothesis 1, on virtually every comparison shown in Tables 2 and 3 Black SPH Residents revealed characteristics unfavorable to smoking cessation, and 24 of the 45 comparisons show differences in proportions significantly different from zero ($\alpha = .05$). These bivariate relationships remain when adjusted for education. Among Black SPH Residents, increased education is significantly associated only with smoking initiation after age 17, increased recognition of risk associated with current smoking level, and awareness of sources of support for quitting (see Table 4). Higher education is negatively associated with plans to quit (not significant) and worry about getting lung cancer.

Regarding Hypothesis 2, there is much less consistency in the comparisons shown in Tables 2 and 3 between Other Blacks and Whites. Compared with Whites, Other Blacks started smoking at an older age; smoke fewer cigarettes per day; have a stronger desire to quit and perceive less effort needed to quit; are more likely to expect to get lung cancer and to be very concerned about the effect of smoking on health in general; and are more likely to consider health concerns, restrictions at work, the expense of cigarettes, and showing that they are not hooked on cigarettes as very important reasons to quit.

On the other hand, compared with Whites, Other Blacks are more likely to believe the risk of lung cancer is the same for smokers and nonsmokers, are more concerned about feeling uncomfortable around smokers and being unable to deal with stress if they quit, and are less likely to know where to get help in quitting. These relationships remain after adjusting for education. However, the adjusted odds ratios and the percentages by education level shown in Table 5 indicate that, in general, educational differences are more consistent than racial differences and that as education increases, racial differences diminish. Thus differences by race are most apparent among the least educated. Education is negatively related to age of smoking onset, amount smoked, and perceived difficulties with quitting; it is positively related to plans to quit, beliefs that smoking is related to lung cancer and that lung cancer would have serious consequences, and knowledge of where to go for help in quitting. Those with more education are also more likely to live in environments that are supportive of cessation and that have fewer smokers.

Discussion

Study Limitations

This was a mixed-mode study, with face-to-face interviews with the SPH sample and telephone interviews with the others; the questionnaires, however, were identical, and there are no indications of method effects in responses to either mode, although the telephone response rate was lower (73% vs 96%). The uniqueness of the SPH sample may limit generalizability, but this population is a growing segment in many urban areas and hence is important to study. Finally, in this study we exam-

TABLE 2—Smoking Habits, Motivation to Quit, Health Belief Model Variables,[16–18] Smoking Prevalence, Social Norms and Social Support for Quitting, and Knowledge of Help Available by Three Study groups

	Black SPH Residents, % (n = 246)	Other Blacks, % (n = 117)	Whites, % (n = 496)	% Difference (95% CI)	
				Black SPH Residents vs Other Blacks	Other Blacks vs Whites
Smoking habits					
Started to smoke at age 17 or younger	68	46	56	22 (10, 34)	−10 (−21, 1)
Smoke more than 10 cigarettes per day	51	35	70	16 (4, 27)	−35 (−45, 25)
Mean no. cigarettes per day (SD)	15.6 (11.8)	12.2 (8.9)	18.5 (10.8)	3.4 (2.3, 4.5)	−6.34 (−8.4, −4.2)
Smoke menthol cigarettes	96	85	49	11 (5, 17)	36 (26, 46)
Motivation to quit smoking					
Ever tried to quit	56	66	67	−10 (−21, 1)	−1 (−10, 8)
Have desire to quit					
None or weak	54	32	35	22 (7, 29)	−3 (−13, 6)
Moderate	27	40	48	−13 (−24, −2)	−8 (−18, 3)
Strong	18	28	17	−10 (−20, 0)	11 (1, 20)
Plan to ever quit	72	90	88	−18 (−26, 9)	2 (−6, 7)
Plan to quit within 12 months	33	58	54	−15 (−36, −14)	4 (−5, 14)
Health Belief Model variables ("Agree" or "strongly agree" with each of the following statements)					
Perceived severity					
These days, there is a good chance that a person who is diagnosed as having lung cancer can be cured.	47	51	43	−4 (−15, 7)	8 (−2, 18)
If I were to get lung cancer, my whole life would change.	79	78	88	1 (−8, 10)	−10 (−17, −3)
If I were to get lung cancer, it would place a great burden on my family or friends.	81	80	87	1 (−8, 10)	−7 (−14, 3)
Perceived susceptibility					
I do not expect to get lung cancer during my life.	41	33	48	8 (−1, 15)	15 (5, 25)
My chances of avoiding lung cancer are better than other people my age.	31	16	21	15 (5, 25)	−5 (−13, 3)
A person has to smoke a lot more than I do to have a serious risk of getting lung cancer.	35	21	17	14 (4, 24)	4 (−14, 12)
Perceived risk					
The chances of getting lung cancer are the same whether or not you smoke.	57	46	27	11 (1, 21)	18 (9, 27)
Concern with risk					
I worry about getting lung cancer more than other people.	30	39	39	−9 (−20, 2)	0
I seldom worry about getting lung cancer.	61	63	60	−2 (−13, 9)	3 (−7, 13)
(The following variables were determined by respondents' answers to questions)					
Benefits of quitting					
Perceive a decrease in likelihood of getting lung cancer after quitting smoking[a]	54	55	60	1 (−4, 3)	−5 (−8, 4)
Other health concerns					
Very concerned about the effects of smoking cigarettes on health	41	46	38	−5 (−17, 6)	8 (−3, 8)
Cost of quitting					
"Very concerned" with the following problems one may experience if quitting:[b]					
Having withdrawal symptoms	39	43	30	−4 (−15, 7)	13 (4, 22)
Feeling uncomfortable around people who smoke	30	23	16	7 (−3, 17)	7 (−2, 16)

Continued

TABLE 2—Continued

	Black SPH Residents, % (n = 246)	Other Blacks, % (n = 117)	Whites, % (n = 496)	% Difference (95% CI): Black SPH Residents vs Other Blacks	% Difference (95% CI): Other Blacks vs Whites
Being unable to deal with stress	35	24	22	11 (1, 21)	2 (−6, 10)
Gaining weight	44	57	54	−13 (−24, −2)	3 (−7, 13)
Perceive effort needed to stop smoking above 5 on a 10-point scale	53	39	58	14 (3, 26)	−19 (−30, −9)
Smoking prevalence					
Most or all relatives smoke	41	37	31	4 (−7, 15)	6 (−3, 15)
Most or all friends smoke	61	46	40	15 (5, 25)	6 (−3, 15)
Social norm for quitting					
People closest to you want you to quit					
Very much	29	57	54	−28 (−39, −16)	3 (−7, 14)
Somewhat	31	22	24	9 (−1, 19)	−2 (−11, 7)
Very little or not at all	40	20	22	20 (7, 30)	−2 (−11, 7)
Expected support for quitting					
Could expect support from					
Smokers in household	43	47	52	−4 (−15, 7)	−5 (−15, 5)
Nonsmokers in household	72	80	73	−8 (−18, 2)	7 (−2, 16)
Relatives living outside household who smoke	56	80	74	−24 (−35, −13)	6 (−3, 15)
Relatives living outside household who do not smoke	69	89	89	−20 (−32, −10)	0
Friends who smoke	52	73	83	−21 (−32, −10)	−10 (−18, −2)
Friends who do not smoke	73	86	91	−13 (−22, −4)	−5 (−11, 1)
Would know where to get help to quit if needed	38	47	71	−9 (−16, −1)	−24 (−34, −14)

Note. SPH = subsidized public housing; CI = confidence interval.
[a]Perception of a decrease was determined by a positive difference between a respondent's ratings on a scale of 1 to 10 of her likelihood of getting lung cancer if she quit and if she continued to smoke.
[b]Asked only of respondents with at least a weak desire or any plans to quit (ns = 225, 110, 476).

TABLE 3—Reasons for Wanting to Quit,[a] by Three Study Groups

"Very Important"	Black SPH Residents, % (n = 179)	Other Blacks, % (n = 105)	Whites, % (n = 440)	% Difference (95% CI): Black SPH Residents vs Other Blacks	% Difference (95% CI): Other Blacks vs Whites
To help current health problems[b]	47	53	32	−6 (−18, 6)	21 (11, 31)
To avoid future health problems	70	84	74	−14 (−23, −5)	10 (2, 18)
Because of health problems of people with whom they live[b]	50	43	36	7 (−5, 19)	7 (−3, 17)
To feel better[b]	65	75	62	−10 (−21, 1)	13 (3, 23)
Because of restrictions at work[b]	6	22	9	−16 (−24, −8)	13 (6, 20)
Because of the expense of smoking	52	55	37	−3 (−15, 9)	18 (3, 28)
To show that they are not hooked on cigarettes	26	38	17	−12 (−23, −1)	21 (12, 30)

Note. SPH = subsidized public housing; CI = confidence interval.
[a]Asked only of respondents who have any desire to quit and who plan to quit someday.
[b]Cases for which question was not applicable are included as "not important" in the total n.

ined characteristics of smokers only. Different patterns might be found if former smokers and nonsmokers were also included.

Implications

Despite these limitations, the present study does have implications for smoking cessation interventions. First, the message that smoking increases the risk of lung cancer has not reached or been accepted by a large proportion of young Black female smokers with less education, although they are concerned about the effect of cigarettes on their health in general. Compared with White or better-educated females, Black females living in SPH and those with less education tend to be very concerned about their current health problems or those of people close to them, the expense of cigarettes, and being addicted to cigarettes. Smoking cessation messages may gain increased attention by emphasizing these concerns. Sec-

TABLE 4—Comparison of Black Subsidized Public Housing Residents (SPH) with Other Blacks (OB): Unadjusted and Education-Adjusted Odds Ratios for Selected Dependent Variables[a]

	Education, %						Unadjusted Odds Ratio, SPH/OB (95% CI)	Adjusted Odds Ratio[b]	
	Low		Mid		High			SPH/OB (95% CI)	High/Low Education[c] (95% CI)
	SPH	OB	SPH	OB	SPH	OB			
Started smoking at age 17 or younger	77	64	66	60	57	34	2.5 (1.6, 4.0)	2.0 (1.5, 2.5)	.6 (.3, .9)
Have strong desire to quit	24	32	11	23	18	29	.6 (.3, 10)	.5 (.02, 1.1)	.8 (.5, 1.2)
Plan to quit	80	86	69	80	63	95	.3 (.2, .6)	.3 (.01, .9)	.8 (.5, 1.1)
Believe one must smoke more than they do to have lung cancer risk	50	14	30	13	19	28	2.0 (1.2, 3.4)	1.7 (1.1, 2.2)	.7 (.4, .9)
Believe the likelihood of lung cancer is same for smokers and nonsmokers	59	62	51	45	61	42	1.6 (1.0, 2.4)	1.5 (1.0, 1.9)	.9 (.6, 1.2)
Worry about lung cancer more than other people	44	46	23	50	16	32	.7 (.4, 1.1)	.5 (.05, 1.0)	.6 (.3, .9)
Would know where to go for help	31	18	44	58	43	51	.7 (.4, 1.1)	.9 (.3, 1.4)	1.4 (1, 1.7)

[a]Omitted are dependent variables for which the comparison presented in Tables 2 and 3 is not significantly altered by education level.
[b]Adjusted using a simultaneous logistic regression model that included study group and education level.
[c]Education was treated as a covariate with three levels (low = less than high school, mid = high school, high = more than high school). The odds ratio indicates the odds for two adjacent levels (high to mid, mid to low).

ond, interventions must address the high smoking prevalence and weak norms and support for quitting in the social environment of low-socioeconomic status female smokers and the likely side effects of cessation. Third, help with cessation must be more readily available to and more heavily promoted in this population, whose members are unaware of the help available. Increased visibility of smoking cessation programs may enhance both their use and smokers' perception of the social desirability of quitting. Finally, those who target programs toward poorly educated young Black female smokers must realize that this group includes women with intensified needs and barriers—exemplified by the public housing residents who are the main focus of this study—who need special attention. □

Acknowledgments

This research was supported by Grant CA42760 from the National Cancer Institute. The authors are grateful to Drs Thomas Cook, Kathleen Crittenden, Frederick Kviz, and Patricia Langenberg for helpful suggestions on earlier drafts.

References

1. Schramm WF. WIC prenatal participation and its relationship to newborn Medicaid costs in Missouri: a cost/benefit analysis. *Am J Public Health.* 1985;75:851–857.
2. Warnecke R, Graham S, Rosenthal S, Manfredi C. Social and psychological correlates of smoking behavior among black women. *J Health Soc Behav.* 1978;19:397–410.
3. US Department of Health and Human Services. *Report of the Secretary's Task Force on Black and Minority Health. Volume III: Cancer.* Washington, DC: US Government Printing Office; 1986. DHHS publication 86-0017.
4. Fiore MJ, Novotny TE, Pierce JP, Hatziandreau EJ, Patel KM, Davis RM. Trends in cigarette smoking in the United States: the changing influence of gender and race. *JAMA.* 1989;261:49–55.
5. Pierce JP, Fiore MC, Novotny TE, Hatziandreu EJ, Davis RM. Trends in cigarette smoking in the United States: educational differences are increasing. *JAMA.* 1989;261:56–60.
6. Pierce JP, Fiore MC, Novotny TE, Hatziandreu EJ, Davis RM. Trends in cigarette smoking in the United States: projections to the Year 2000. *JAMA.* 1989;261:61–65.
7. US Department of Health and Human Services. *The Health Consequences of Smoking: Nicotine Addiction.* Report of the Surgeon General. DHHS; 1988. NIH publication 88-8406. Centers for Disease Control, Rockville, Md: Office on Smoking and Health.
8. Novotny TE, Warner KE, Kendrick JS, Remington PL. Smoking by Blacks and Whites: socioeconomic and demographic differences. *Am J Public Health.* 1988; 78:1187–1189.
9. Leffall LD. Program priorities for the future. In: Mettlin C, Murphy GP, eds. *Cancer among Black Populations.* New York, NY: Alan R. Liss; 1981:253–263.
10. Cardwell JJ, Collier WV. Racial differences in cancer awareness. *Urban Health.* October 1981:29–32.
11. Auletta K. *The Underclass.* New York: Random House; 1982.
12. Glasgow DG. *The Black Underclass: Poverty, Unemployment, and Entrapment of Ghetto Youth.* San Francisco, Calif: Jossey-Bass; 1980.
13. Wilson WJ. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* Chicago, Ill: University of Chicago Press; 1987.
14. Warnecke RB, Flay BR, Kviz FJ, et al. Characteristics of participants in a televised smoking cessation intervention. *Prev Med.* 1991;20:389–403.
15. Orleans CT, Schoenbach V, Salmon M, et al. A survey of smoking and quitting patterns among Black Americans. *Am J Public Health.* 1989;79:176–180.
16. Becker MH, Haefner DP, Kasl SV, Kirscht JP, Maiman LA, Rosenstock IM. Selected psychosocial models and correlates of health-related behaviours. *Med Care.* 1977;15 (May suppl):27–46.
17. Rosenstock IM. The Health Belief Model and preventive health behavior. In: Becker MH, ed. *The Health Belief Model and Personal Health Behavior.* Thorofare, NJ: Charles B. Slack; 1974:27–59.
18. Kviz FJ, Crittenden KS, Belzer LJ, Warnecke RB. Psychosocial factors in enrollment in a televised smoking cessation program. *Health Educ Q.* In press.
19. Minium EW. *Statistical Reasoning in Psychology and Education.* 2nd ed. New York, NY: John Wiley & Sons Inc; 1978.
20. Hosmer DW, Lemeshow S. *Applied Logistic Regression.* New York, NY: John Wiley & Sons Inc; 1989.

TABLE 5—Comparison of Other Blacks (OB) with Whites (W): Unadjusted and Education-Adjusted Odds Ratios for Selected Dependent Variables[a]

	Education %							Adjusted Odds Ratio[b]	
	Low		Mid		High		Unadjusted Odds Ratio, OB (95% CI)	Other Blacks/Whites (95% CI)	High/Low Education[c] (95% CI)
	OB	W	OB	W	OB	W			
Started smoking at age 17 or younger	64	76	60	61	34	51	.7 (.4, .9)	.6 (.4, .9)	.4 (.2, .7)
Smoke >10 cigarettes per day	36	84	43	76	31	63	.2 (.2, .4)	.2 (.2, .7)	.6 (.4, .9)
Plan to quit	86	73	80	87	95	92	1.1 (.6, 2.1)	1.2 (.6, 1.9)	2.0 (1.6, 2.4)
Plan to quit within 12 months	55	30	50	51	63	60	1.2 (.8, 1.8)	1.2 (.8, 1.6)	1.6 (1.3, 1.8)
Agree or strongly agree that "if I were to get lung cancer my whole life would change"	60	81	83	86	82	90	.5 (.3, .8)	.5 (0, 1.0)	1.5 (1.2, 1.8)
Agree or strongly agree that "the chances of getting lung cancer are the same whether or not you smoke"	62	49	45	34	42	19	2.3 (1.5, 3.5)	2.3 (1.9, 2.7)	.5 (.3, .8)
Very concerned about effect of smoking on health	59	35	47	40	42	38	1.4 (.9, 20)	2.3 (1.9, 2.7)	.5 (.3, .8)
Concerned about withdrawal symptoms if they quit	67	40	44	37	34	24	1.7 (1.1, 2.6)	1.6 (1.0, 2.4)	.5 (.3, .9)
Concerned about feeling uncomfortable around smokers if they quit	42	24	26	16	19	16	1.7 (1.0, 2.7)	1.6 (1.1, 2.1)	.8 (.3, .9)
Concerned about gaining weight if they quit	75	63	67	54	48	52	1.2 (.8, 1.8)	1.2 (.7, 1.6)	.8 (.5, 1.0)
Concerned about being unable to deal with stress if they quit	55	14	30	23	13	24	1.1 (.7, 1.8)	1.1 (.6, 1.6)	.7 (.6, 1.1)
People close to respondents very much want them to quit	41	46	60	50	61	58	1.1 (.8, 1.7)	1.2 (.8, 1.6)	1.3 (1.1, 1.6)
Most or all relatives smoke	46	35	37	39	28	18	1.3 (.9, 2.0)	1.3 (.8, 1.7)	.5 (.3, .8)
Most or all friends smoke	50	54	37	46	40	32	1.1 (.7, 1.6)	1.0 (.6, 1.5)	.7 (.4, .9)
Expect support from friends who do not smoke	68	84	90	92	91	92	.6 (.3, 1.1)	.6 (.1, 1.2)	1.6 (1.2, 2.0)
Expect support from friends who smoke	46	81	67	82	86	84	.6 (.4, .9)	.6 (.1, 1.1)	1.5 (1.2, 1.8)
Would know where to get help to quit	18	53	58	68	52	76	.4 (.2, .6)	.4 (.1, .5)	1.7 (1.4, 1.9)
Important reasons to quit									
Current health problems	79	41	67	54	40	26	2.7 (1.7, 4.6)	2.5 (2.0, 2.9)	.6 (.3, .8)
Health problems of people with whom they live	68	52	42	40	36	31	1.3 (.8, 2.1)	1.3 (.8, 1.7)	.6 (.5, .9)
To feel better	90	63	75	70	71	57	1.9 (1.1, 2.9)	1.8 (1.3, 2.3)	.7 (.4, 1.0)
Restrictions at work	42	15	17	9	18	8	2.9 (1.6, 5.1)	2.8 (2.2, 3.4)	.7 (.3, 1.1)
Expense of smoking	84	56	58	44	45	29	2.1 (1.4, 3.3)	2.1 (1.7, 2.5)	.5 (.2, .8)
To show they are not hooked	68	15	25	20	34	15	3.0 (1.9, 4.8)	2.9 (2.5, 3.4)	.7 (.4, 1.0)

[a]Omitted are dependent variables for which the comparison presented in Tables 2 and 3 is not significantly altered by education level.
[b]Adjusted using a simultaneous logistic regression model that included study group and education level.
[c]Education was treated as a covariate with three levels (low = less than high school, mid = high school, high = more than high school). The odds ratio indicates the odds for two adjacent levels (high to mid, mid to low).

Smoking Initiation and Cessation in Relation to Body Fat Distribution Based on Data from a Study of Swedish Women

Lauren Lissner, PhD, Calle Bengtsson, MD, Leif Lapidus, MD, Cecilia Björkelund, MD

ABSTRACT

In a representative sample of Swedish women, smokers were significantly less obese than nonsmokers. However, a smoker was likely to have significantly more upper-body fat than a nonsmoker of similar body mass index. Women who quit smoking experienced less upper-body fat deposition than would be expected by their accompanying weight gain, suggesting that weight gained as a consequence of smoking cessation is not preferentially deposited in the region associated with increased cardiovascular risk. (*Am J Public Health.* 1992;82:273–275)

Introduction

A recent survey showed that 39% of female cigarette smokers viewed smoking as a tool for weight control.[1] Weight gain and increased appetite were cited as important reasons for relapse following attempts to quit smoking. In agreement with this popular notion, epidemiological studies have demonstrated clearly that smokers weigh less than nonsmokers[2,3] and that smoking cessation leads to weight gain among both men[3–5] and women.[4–6] In two studies involving male subjects, associations have been observed between cigarette smoking and centralized body fat distribution.[7,8] Furthermore, in one of these latter studies, smoking cessation was associated with a smaller degree of upper-body fat deposition than would be anticipated from the accompanying weight gain.[7] Less is known, however, about the effects of smoking initiation and cessation on body fat distribution in women.

Thus, the aim of this paper is twofold: to describe cross-sectional associations between smoking and body fat distribution in a representative sample of Swedish women; and to examine changes in fat distribution that occur as a consequence of smoking initiation and cessation. If the findings in women are similar to those in men, this would confirm that gains in body weight following smoking cessation do not follow the androgenic fat distribution pattern associated with increased cardiovascular risk in women[9–12] and in men.[11–13]

Subjects and Methods

This analysis is based on data from the Prospective Population Study of Women in Gothenburg, Sweden. This is a representative sample of female Gothenburg residents born in 1908, 1914, 1918, 1922, and 1930. The analysis presented here uses data from the first two examinations, which took place in 1968/69 and in 1974/75. Details of the sampling, participation rates, and examination protocol have been given elsewhere.[6,9,14,15]

Anthropometric measurements were taken at the same time of day and year at both examinations. A subject's body mass index (BMI) ($weight_{kg}/height_m^2$) was used as an indicator of overall adiposity, while waist-hip ratio (waist circumference/hip circumference) reflected the distribution of her body fat, with high values indicating upper-body obesity. Changes in BMI and waist-hip ratio were calculated as 1974/75 values minus 1968/69 values. Smoking history was ascertained by means of standardized questions at each examination. Subjects who had never smoked or who had not smoked for at least a year were classified as nonsmokers. Women who reported quitting at any time during the previous 12 months were excluded.

In the cross-sectional portion of this investigation, analysis of variance was used to compare BMI and waist-hip ratio in current smokers with those in nonsmokers, using 1974/75 data. All cross-sectional findings were replicated using the 1968/69 measurements not shown. Similarly, the longitudinal analysis used data from both examinations to compare continuing smokers with women who quit smoking, and nonsmokers with women who started smoking. After controlling for age and (when appropriate) BMI, adjusted least square means were computed and reported here with their standard errors.

Lauren Lissner, Calle Bengtsson, and Cecilia Björkelund are with the Department of Primary Health Care and Leif Lapidus is with the Department of Medicine at Sahlgrenska Hospital at the University of Gothenburg in Gothenburg, Sweden.

Requests for reprints should be sent to Lauren Lissner, PhD, Department of Primary Health Care, Gothenburg University, Redbergsvagen 6, 41665 Gothenburg, Sweden.

This paper was submitted to the journal September 10, 1990, and accepted with revisions June 10, 1991.

TABLE 1—Comparisons of Smokers and Nonsmokers, 1974/75: Age-adjusted (top); Adjusted for Age and Body Mass Index (bottom).

	Smokers		Nonsmokers		Smokers vs Nonsmokers
	(n = 486)	SE	(n = 805)	SE	*P*
BMI[a]	23.8	(0.18)	25.1	(0.14)	.0001
waist-hip ratio	0.793	(0.003)	0.782	(0.002)	.003
BMI[a]-adjusted waist-hip ratio	0.798	(0.003)	0.779	(0.002)	.0001

Note. Least square means are shown with standard errors.
[a]BMI = body mass index.

TABLE 2—Effects of Smoking Initiation and Cessation between Examinations 1 and 2 on Body Mass Index and Waist-Hip Ratio: Age-adjusted (top); Adjusted for Age and Change in Body Mass Index (bottom).

	Initiation			Cessation		
	Never Smoked	Started		Continue Smoking	Quit	
Change in	(n = 701)	(n = 26)	*P* diff	(n = 435)	(n = 63)	*P* diff
BMI[a]	0.62 (0.067) ****	−0.37 (0.35)	**	0.50 (0.089) ****	1.4 (0.24) ****	***
Waist-hip ratio	0.048 (0.0016) ****	0.046 (0.0084) ****	n.s.	0.050 (0.003) ****	0.040 (0.0079) ****	n.s.
BMI[a]-adjusted waist-hip ratio	0.048 (0.0015) ****	0.054 (0.0079) ****	n.s.	0.051 (0.0029) ****	0.033 (0.0077) ****	*

Note. For each group, least square means are shown with standard errors and significance for H_0: change = 0. Right columns show significance levels for between-group differences.
[a]BMI = body mass index.
*P < .05, **P < .01, ***P < .001, ****P < .0001. N.S. = not significant.

Results

Cross-sectional Analysis

The rate of cigarette smoking was 41.0% in 1968/69 and 37.6% in 1974/75. Smokers had lower BMIs (P = .0001) and higher waist-hip ratios (P = .003) than nonsmokers, after controlling for age, in agreement with previous findings among this cohort[6,9] (see Table 1). To determine whether smoking was associated with fat patterning independent of generalized obesity, further statistical adjustment was made for BMI. The differences between waist-hip ratios of smokers and nonsmokers were accentuated after controlling for BMI (P = .0001), indicating that a smoker is likely to have a significantly higher proportion of upper- to lower-body fat than a nonsmoker of similar body mass.

Longitudinal Analysis

Between the two examinations, 3.5% of nonsmokers started smoking and 12.7% of smokers quit. The purpose of the second analysis was to evaluate specific effects of smoking initiation and cessation on BMI and waist-hip ratio. To test for an initiation effect, comparisons were made between subjects who did not smoke at either exam and those who began smoking between the two exams. As expected, the women who started smoking tended to lose weight while the nonsmokers gained weight, and the difference between these groups was significant (P = .005). Conversely, a comparison of women who smoked at both exams with those who quit indicated that, although both groups gained weight, the quitters gained significantly more (P = .0006) (see Table 2).

To evaluate effects of smoking initiation and cessation on fat distribution, changes in waist-hip ratio were examined, with and without adjustment for changes in BMI (see Table 2). The initial analysis showed no significant effects of smoking initiation or cessation on waist-hip ratio; all four groups displayed significant increases in this ratio over time (P = .0001). However, after controlling for possible confounding by BMI changes, significant differences in the waist-hip ratio emerged in smokers versus quitters (P = .03). Waist-hip ratio increased by slightly more in smokers (+0.05) than in quitters (+0.03), which is contrary to what one would predict from their respective BMI gains of +0.5 and +1.4. It is concluded that women who stop smoking do not experience the degree of upper-body fat deposition that generally accompanies weight gain.

Discussion

International comparisons of cigarette smoking trends suggest that women are quitting less often and/or starting more often than men worldwide.[16,17] At current rates, it has been projected that the majority of smokers in the United States could be women after the year 1995.[18] Therefore, the observation that smoking cessation resulted in less upper-body fat deposition than would be expected from the concomitant weight gain may be of interest to many female smokers. Upper-body fat distribution in women is related to increased rates of coronary heart disease,[9,11] stroke,[9] diabetes,[12] and unfavorable lipid profiles,[10,12] and it may be a more potent cardiovascular risk factor than generalized obesity.[9-12] In this context, our results suggest that the weight gain that often accompanies smoking cessation need not be perceived as an adverse health effect because weight gained as a consequence of quitting smoking is not preferentially deposited in the region associated with increased cardiovascular risk. Such information, together with the similar findings in males,[7] could provide an incentive to a health-conscious public to modify smoking practices in line with a more advantageous body fat distribution, even if some overall weight gain is involved. □

Acknowledgment

This work was supported by the Swedish Medical Research Council. The authors thank Reubin Andres and Hiroshi Shimokata for useful discussions on this topic.

References

1. Klesges R, Klesges L. Smoking and fear of weight gain. *Int J Eating Disord.* 1988;7:413–419.
2. Albanes D, Jones DY, Micozzo MS, et al. Associations between smoking and body weight in the US population: analysis of NHANESII. *Am J Public Health.* 1987;77:439–444.
3. Gordon T, Kannel WB, Dawber TK, et al. Changes associated with quitting cigarette smoking: the Framingham Study. *Am Heart J.* 1975;90:322–328.
4. Williamson DF, Mandans J, Anda RF, et al. Smoking cessation and severity of weight gains in a national cohort. *N Engl J Med.* 1991;324:739–745.
5. Lund-Larsen P, Tretli S. Changes in smoking habits and body weight after a 3-year period: the Cardiovascular Disease Study in Finnmark. *J Chron Dis.* 1982;35:773–780.
6. Noppa H, Bengtsson C. Obesity in relation to smoking: a population study of women in Göteborg, Sweden. *Prev Med.* 1980;9:534–543.
7. Shimokata H, Muller D, Andres R. Studies in the distribution of body fat, III: effect of cigarette smoking. *JAMA.* 1989;261:1169–1173.
8. Comstock GW, Stone RW. Changes in body weight and subcutaneous fatness related to smoking habits. *Arch Environ Health.* 1972;24:271–276.
9. Lapidus L, Bengtsson C, Larsson B, et al. Distribution of adipose tissue and risk of cardiovascular disease and death: a 12-year follow-up of participants in the population study of women in Gothenburg, Sweden. *Br Med J.* 1984;289:1257–1261.
10. Seidell JC, Cigolini M, Charzewska J, et al. Indicators of fat distribution, serum lipids, and blood pressure in European women born in 1948: the European Fat Distribution Study. *Am J Epidemiol.* 1989;130:53–65.
11. Stokes J, Garrison RK, Kannel WB. The independent contributions of various indices of obesity to the 22-year incidence of coronary heart disease: the Framingham Heart Study. In: Vague J, Björntrp P, Guy-Grand B, Rebuffé-Scrive M, eds. *Metabolic Complications of Human Obesities.* Amsterdam, Holland: Elsevier Science Publishers; 1985:49–57.
12. Haffner SM, Stern MP, Hazuda HP, et al. Do upper-body and centralized adiposity measure different aspects of regional body-fat distribution? Relationship to non-insulin-dependent diabetes mellitus, lipids, and lipoproteins. *Diabetes.* 1987;36:43–51.
13. Larsson B, Svärdsudd K, Welin L, et al. Abdominal adipose tissue distribution, obesity and risk of cardiovascular disease and death: a 13-year follow-up of participants in the study of men born in 1913. *Br Med J.* 1984;288:1401–1404.
14. Bengtsson C, Blohmé G, Hallberg L, et al. The study of women in Gothenburg 1968–1969—a population study: general design, purpose, and sampling results. *Acta Med Scand.* 1973;193:311–318.
15. Bengtsson C, Hallberg L, Hällström T, et al. The population study of women in Göteborg 1974–1975—the second phase of a longitudinal study: general design, purpose and sampling results. *Scand J Soc Med.* 1978;6:49–54.
16. Pierce JP. International comparisons in trends in cigarette smoking prevalence. *Am J Public Health.* 1989;79:152–157.
17. Fiore MC, Novotny TE, Pierce JP, et al. Trends in cigarette smoking in the United States: the changing influence of gender and race. *JAMA.* 1989;261:49–55.
18. Pierce JP, Fiore MC, Novotny TE, et al. Trends in cigarette smoking in the United States: projections to the year 2000. *JAMA.* 1989;261:61–65.

Surveillance of Occupational Lead Exposure in New Jersey: 1986 to 1989

Allison Tepper, PhD

ABSTRACT

Between January 1986 and June 1989, 1916 New Jersey workers were identified through a surveillance system for occupational lead exposure. The average annual proportion of workers with a blood lead level above 2.42 μmol/L was 12%. Industries with the highest proportion of workers with blood lead levels above 2.42 μmol/L were special trade construction (35%) and industries dealing with scrap and waste materials (27%). (*Am J Public Health.* 1992;82:275–277)

Introduction

Public health agencies in several states have described surveillance programs for occupational lead exposure using biological monitoring data.[1–3] The experience in these states provides useful lessons about program design and operation for others beginning to implement lead surveillance activities. Ongoing reporting of findings helps promote the use of surveillance data for targeting and evaluating lead poisoning prevention efforts. This paper examines the results obtained by the New Jersey Department of Health (NJDOH) occupational lead exposure surveillance project between January 1986 and June 1989.

Methods

A passive surveillance system for occupational lead exposure was begun by the NJDOH in October 1985. In-state clinical laboratories are required to report test results and identifying information for adults (aged 16 and above) with blood lead levels above 1.21 μmol/L* (New Jersey Administrative Code 8:44–2.11).

The NJDOH provides educational materials to physicians and reported individuals by mail and conducts telephone interviews to ascertain the source of exposure. When a work-related exposure is

*1 μmol/L = 20.7 μg/dL.

At the time of this study, the author was with the Occupational Health Service, New Jersey Department of Health. She is now with the National Institute for Occupational Safety and Health (NIOSH), Division of Surveillance, Hazard Evaluation and Field Studies.

Requests for reprints should be sent to Allison Tepper, PhD, NIOSH, Division of Surveillance, Hazard Evaluation and Field Studies, 4676 Columbia Parkway, Mail Stop R-10, Cincinnati, OH 45226.

This paper was submitted to the journal September 24, 1990, and accepted with revisions June 19, 1991.

TABLE 1—Reports Received and Workers Identified by the New Jersey Department of Health, by Date of Blood Test: January 1986–June 1989

Test Date	Reports Received n	Reports Received %	New Workers Identified n	New Workers Identified %
1986				
Jan–June	735	10	433	23
July–Dec	519	7	223	12
1987				
Jan–June	411	6	116	6
July–Dec	443	6	108	6
1988				
Jan–June	1020	14	378	20
July–Dec	1651	23	250	13
1989				
Jan–June	2292	32	236	12
Total	7071	100	1916[a]	100[a]

[a]86 workers (4%) were first reported in 1985; the date of first report was missing for 86 workers (4%).

identified, a NJDOH representative visits the workplace or refers the matter to OSHA.

Results

Surveillance Findings

Of 2487 reported individuals, 1916 (77%) were exposed to lead in their New Jersey workplaces. The number of reports increased after the first 2 years, but there was no discernible trend in the number of new workers reported (Table 1).

The distribution of blood lead levels was consistent from year to year. On average, 27% of those reported had an annual peak blood lead level from 1.21 to 1.44 μmol/L; 43%, from 1.45 to 1.92 μmol/L; 19%, from 1.93 to 2.41 μmol/L; 7%, from 2.42 to 2.89 μmol/L; 3%, from 2.90 to 3.37 μmol/L; and 2%, above 3.38 μmol/L.

Nine facilities in four industries (cyclic crudes and intermediates, storage batteries, primary copper, and vitreous china tableware and kitchenware) accounted for 959 workers (51%) (Table 2). Forty-three facilities in 11 other industries accounted for 486 workers (26%), including 210 workers (11 percent) in metal-related industries and 143 workers (8%) in plastics-related industries. One hundred twenty-seven facilities in 77 industries accounted for the remaining 425 workers (23%), including 89 workers (5%) in the construction industry.

Blood lead levels above 2.42 μmol/L were reported for 246 workers (13%) (Table 3). Thirty-three percent of construction workers (84% of whom were special trade contractors) and 26% of wholesale trade workers (82% of whom worked in scrap and waste materials facilities) had blood lead levels above 2.42 μmol/L.

Characteristics of the Surveillance System

Ten in-state and three out-of-state laboratories provided reports to the NJDOH. Blood lead levels below 1.93 μmol/L were reported within an average of 17.9 days, compared with 19.8 days for higher levels ($P = .08$). Although laboratories nearly always provided the workers' names, blood lead levels, and test dates, most reports did not provide age (54%), address (61%), and telephone number (84%).

Discussion

The New Jersey experience demonstrates the utility of laboratory-based surveillance systems for occupational lead exposure. Positive features include clear reporting criteria, small numbers of reporting units, high specificity for detecting cases of interest, and the identification of workplaces that can be targeted for intervention. Although some high-risk industries, such as secondary smelting and battery manufacturing, are common to many states,[3] state specific, laboratory based surveillance systems identify industries important in one state but not in others, such as cyclic crudes and intermediates (which includes organic lead production) in New Jersey. Improvements are needed if these surveillance systems are to provide accurate data for assessing progress toward the Public Health Service objective of eliminating blood lead levels above 2.42 μmol/L by the year 2000.[4]

Failure to capture test results from out-of-state laboratories contributes to underestimation of the number of overexposed workers. A large increase in the number of new cases in New Jersey resulted from the initiation of voluntary reporting by an employer using an out-of-state laboratory. A California survey showed that 47% of lead-using facilities sent their specimens to out-of-state laboratories. Fifty-three percent of reports to the New York heavy metals registry were from out-of-state laboratories.[1] The feasibility of using state licensing authority to cover out-of-state laboratories (this authority is present in New York but not in New Jersey) should be considered by all states interested in lead surveillance. If possible, laboratory certification requirements of the OSHA lead standard should be modified to include provisions for reporting to states. The sensitivity of sur-

TABLE 2—Workers and Facilities Reported to the New Jersey Department of Health, by Standard Industrial Classification (SIC) Code: January 1986–June 1989

SIC	Workers n	Workers %	Facilities n	Facilities %
2865—Organic chemicals	324	17.3	5	2.8
3691—Storage batteries	263	14.1	2	1.1
3331—Primary nonferrous metals	213	11.4	1[a]	.6
3262—Vitreous china food utensils	159	8.5	1	.6
3079—Miscellaneous plastics products	93	5.0	7	3.9
3341—Secondary smelting of nonferrous metals	59	3.2	4	2.2
2821—Plastics materials and resins	50	2.7	2	1.1
3312—Blast furnaces and steel mills	50	2.7	2	1.1
3356—Nonferrous rolling and drawing	48	2.6	4	2.2
5093—Scrap and waste materials	41	2.2	11	6.1
3679—Electronic components	37	2.0	1	0.6
3362—Brass foundry	32	1.7	4	2.2
2816—Inorganic pigment	28	1.5	5	2.8
3999—Miscellaneous manufacturing	27	1.4	1	0.6
3443—Fabricated plate work	21	1.1	2	1.1
All others[b]	425	22.7	127	70.9
Total	1870[c]	100.0	179[c]	100.0

[a]This facility is no longer operating.
[b]Includes 77 four-digit SIC codes, each with less than 20 workers.
[c]Forty-six workers and 21 facilities with unknown SIC codes are excluded from the table.

veillance systems would be improved by capturing all workers with workplace biological monitoring programs, but would still be limited by the lack of monitoring in a large percentage of lead-using workplaces.[5]

The absence of complete identifying information on many reports hinders follow-up. This problem could be ameliorated by the strengthening of reporting regulations and enforcement policies. Alternate sources of information for surveillance also should be considered. For example, new reporting requirements in New Jersey allow the state to obtain information directly from physicians (New Jersey Administrative Code 8:57–3.2).

Workers in New Jersey are being exposed to lead at levels above limits in the OSHA lead standard. The proportion of New Jersey workers with the highest blood lead levels (above 2.42 μmol/L) was 12%, which is consistent with the findings from California.[2] In the absence of accurate denominator data, however, the true prevalence of overexposure among all lead-exposed workers cannot be determined. Moreover, the distribution of blood lead levels is distorted by the absence of information about blood lead levels below the reporting level.

The most severe lead exposure problems in New Jersey appear to be in the construction and scrap metal industries. In several reports of outbreaks of lead poisoning, exclusion of the construction industry from the OSHA lead standard has been recognized as a problem.[6,7] Additionally, intermittent lead exposure and a mobile work force (particularly in construction) complicate the development of adequate health and safety programs. Finally, many workers in these industries are employed in small businesses, where health and safety programs are less common.[8] In New Jersey, 92% of employers in special trade construction and 86% of employers in scrap metal industries employ fewer than 20 workers. In contrast, only 62% of employers in the manufacturing sector are in this category.[9] □

TABLE 3—Workers Reported to the New Jersey Department of Health with Blood Lead Levels above 2.42 μmol/L, by Standard Industrial Classification (SIC) Code: January 1986–June 1989

SIC	Workers		
		Blood Lead > 2.42 μmol/L	
	Total n	n	%
Construction (15–17)	89	33	37.1
17—Special trade contractors	75	26	34.7
Manufacturing (20–39)	1646	192	11.7
28—Chemicals and allied products	475	58	12.2
30—Rubber & miscellaneous plastics products	98	13	13.3
32—Stone, clay, & glass products	172	12	7.0
33—Primary metal industries	483	62	12.8
34—Fabricated metal products	43	4	9.3
36—Electronic & other electric equipment	312	32	10.3
39—Miscellaneous manufacturing	33	6	18.2
Transportation (40–49)	9	1	11.1
Wholesale Trade (50–51)	50	13	26.0
50—Durable goods	41	11	26.8
Services (70–79)	44	4	9.1
75—Auto repair, services, & parking	20	0	0.0
Public Adminstration (90–97)	32	3	9.8
Total	1870	246	13.2

Note. Date are presented for each major SIC division and for two-digit SIC codes with 20 or more reported workers.

Acknowledgments

The author thanks Dr Barbara Gerwel, Nan Long, and Marion Pierson for their tireless efforts to maintain the blood lead reporting system.

References

1. Baser ME, Marion D. A statewide case registry for surveillance of occupational heavy metals absorption. *Am J Public Health.* 1990;80:162–164.
2. Maizlish N, Rudolph L, Sutton P, Jones JR, Kizer KW. Elevated blood lead in California adults, 1987: results of a statewide surveillance program based on laboratory reports. *Am J Public Health.* 1990;80:931–934.
3. Rudolph L, Maizlish N, Tepper A, et al. Surveillance for occupational lead exposure—United States, 1987. *MMWR.* 1989;38:642–644.
4. *Healthy People 2000.* Washington, DC: US Dept of Health and Human Services; 1989. Publication PHS 91-50212.
5. Rudolph L, Sharp DS, Samuels S, Perkins C, Rosenberg J. Environmental and biological monitoring for lead exposure in California workplaces. *Am J Public Health.* 1990;80:921–925.
6. Himmelstein J, Wolfson M, Pransky G, Morse D, Ross A, Gill J. Lead poisoning in bridge demolition workers—Massachusetts. *MMWR.* 1989;38:687–694.
7. Stockbridge H, Daniell W. Lead poisoning among bricklayers—Washington State. *MMWR.* 1991;40:169–171.
8. Pedersen DH, Sieber WK. *National Occupational Exposure Survey: Analysis of Management Interview Responses.* Cincinnati, Ohio: National Institute for Occupational Safety and Health; 1988. US Dept of Health and Human Services. Publication NIOSH 89-103.
9. US Bureau of the Census. *County Business Patterns, 1986. New Jersey.* Washington, DC: US Government Printing Office; 1988. Publication CBP-86-32.

Inaccuracies in Certification of Nonmelanoma Skin Cancer Deaths

Martin A. Weinstock, MD, PhD, Hendrik A. Bogaars, MD, Michelle Ashley, MD, Virginia Litle, MD, Elise Bilodeau, MD, and Steven Kimmel, MD

ABSTRACT

Nonmelanoma skin cancer is the most common cancer site in the United States, yet mortality from this cause is poorly understood. We sought medical records of the 116 reported deaths during 1979 through 1987 from this cause (*International Classification of Diseases,* 9th version [ICD-9], code 173) among Rhode Island residents to evaluate the accuracy of the reported cause of death. Of the 110 cases for which the cause of death could be classified as correct or incorrect, 59 (54%) were misclassified, 49 (83%) of which were mucous-membrane, squamous-cell carcinomas of the head and neck. For most of these, the written death certificate diagnosis was squamous-cell carcinoma of the head and neck, which was coded 173.4. Other problematic diagnoses were cancer of the head and neck and malignant fibrous histiocytoma. In response to a mailed survey, most health departments replied that squamous-cell carcinoma of the head and neck was coded under rubric 173 and malignant fibrous histiocytoma was coded under rubric 171, but there was no unanimity. The misclassification of other causes of death to ICD-9 rubric 173 is substantial. The vast majority were coded to rubric 173.4 and were due to a small number of diagnoses that are recognizable on examination of the death certificate. (*Am J Public Health.* 1992; 82:278–281)

Introduction

Nonmelanoma skin cancer (NMSC) is by far the most commonly diagnosed malignancy in the United States and is increasing rapidly in incidence.[1–5] Mortality from this cause has been poorly studied, however. Mortality rates based on death certificate data indicate that from the 1950s to the 1970s, there was a substantial decline in death due to NMSC.[6] However, these rates may be inaccurate.[6–8] We reviewed reported NMSC deaths (a) to estimate the accuracy of death certification, (b) to evaluate secular changes in accuracy, (c) to determine the reasons for inaccuracy, (d) to assess implications for the interpretation of national NMSC mortality rate statistics, and (e) to suggest improvements.

Methods

We reviewed certificates and sought medical records for all deaths among Rhode Island residents during 1979 through 1987 for which the underlying cause of death was coded as NMSC (*International Classification of Diseases,* 9th version [ICD-9], code 173). Appendix A describes our criteria for determining whether the cause of death code was correct.

Some of the underlying causes of death written on the death certificates by the physicians certifying the deaths were potentially ambiguous with respect to appropriate ICD-9 code. To aid in distinguishing problems of coding practice from problems in coding policy, we sent a short questionnaire in September 1990 to 53 health departments (50 states, New York City, the District of Columbia, and Puerto Rico). For six diagnoses, we asked "Which ICD-9 code would you use to classify a death certificate with the following underlying cause listed on the death certificate?"

We use the term "accuracy" herein to refer to the proportion of certificates correctly classified under ICD-9 code 173 among all evaluable certificates classified under that rubric. This is the "confirmation rate,"[8] that is, the (positive) predictive value of classification under this rubric if our evaluation of the medical records is taken as the standard of reference.

P values are two-tailed and were calculated using the Student's *t* statistic or linear regression.

Results

During 1979 through 1987, 116 deaths (including 4 out-of-state deaths) among Rhode Island residents were classified under ICD-9 code 173. Six deaths were excluded (see Appendix B). Fifty-nine (54%) of the 110 remaining deaths were incorrectly classifie[illegible]MSC. The mean age at death of the incorrectly classified group was lower (63.3 years) than the correctly classified group (74.5 years; $P < .00001$). There was a significant increase in misclassification over the 9-year period of study ($P = .02$). Autopsies were performed in only four cases, one of which (a noncutaneous, malignant fibrous histiocytoma) was misclassified. The incorrectly classified deaths did not differ from the correctly classified deaths with respect to gender or whether death occurred in a hospital (Table 1).

Table 2 displays the cause of death determined by examination of medical records of misclassified deaths. Striking trends were noted on comparison of the cause actually written on the death certificate by the certifying physician with the correctness of classification (see Table 3). If the problem categories of squamous-cell cancer or carcinoma of the neck or of the head and neck, cancer or carcinoma of the neck or of the head and neck, and malignant fibrous histiocytoma are eliminated, the accuracy of classification increases

The authors are with the Dermatoepidemiology Unit, Department of Medicine, Veterans Administration Medical Center, Roger Williams Medical Center, and Brown University.

Requests for reprints should be sent to Martin A. Weinstock, MD, PhD, Dermatoepidemiology Unit 111, VA Medical Center, 830 Chalkstone Avenue, Providence, RI 02908.

This paper was submitted to the journal February 1, 1991, and accepted with revisions June 10, 1991.

from 46% to 88%, but only five (10%) of the correctly classified deaths are removed.

We received responses from 43 (81%) of the health departments to the single mailing of the questionnaire regarding actual coding practices. Two did not code cause of death themselves. There was no unanimity in coding for the three problematic categories described previously (squamous-cell cancer of the neck or of the head and neck, cancer of the neck or of the head and neck, and malignant fibrous histiocytoma), although for each the majority were consistent with the National Center for Health Statistics (NCHS) recommendations (see Table 4). Some states described inconsistencies in their coding of these diagnoses.

Fifty-five (50%) of the 110 deaths were classified under ICD 173.4 (scalp and neck). These included 48 (98%) of the mucosal carcinomas, the 1 lip carcinoma, and 6 (12%) of the correctly classified NMSC deaths.

TABLE 1—Proportion of ICD 173 Deaths Inaccurately Classified by Age, Gender, Race, Year of Death, and Location of Death: Rhode Island Residents, 1979–1987

	No. Misclassified	% Misclassified
Age (y)		
0–54	13	72
55–64	25	83
65–74	10	43
75–84	6	29
85+	5	28
Gender		
Male	43	57
Female	16	47
Race		
White	55	53
Black	4	67
Year of death		
1979–1981	8	35
1982–1984	24	57
1985–1987	27	60
Location of death		
In hospital	35	54
Out of hospital	24	53

Note. ICD = *International Classification of Diseases.*

Discussion

We noted inaccurate classification of cause for over half of the reported NMSC deaths among Rhode Island residents. The vast majority of the misclassified deaths were due to mucosal squamous-cell carcinoma of the head and neck region. The vast majority of these could be distinguished from true NMSC deaths by examination of the death certificate. Coding of certain diagnoses has been inconsistent among health departments. The policy of coding squamous-cell carcinoma of the head and neck with NMSC (ICD-9 code 173) accounted for more than half of the misclassified deaths. Most (83%) of the misclassified deaths were coded under ICD 173.4.

Our study was subject to certain limitations. Postmortem examinations were rarely performed, although results of histopathologic examinations of relevant surgical specimens were available for almost all deaths. The documentation available in the medical records was frequently suboptimal. We adhered to specific criteria that required some evidence that the NMSC contributed to death; stricter criteria would have resulted in a higher proportion of unclassifiable cases, but our estimates of accuracy then might be biased since these unclassifiable cases would have been primarily individuals with advanced stages of NMSC and no documentation of the events immediately leading to death.

TABLE 2—Cause of Death Determined by Review of Medical Records among Deaths Misclassified as Nonmelanoma Skin Cancer: Rhode Island Residents, 1979–1987

Cause	No.	%
Squamous cell carcinoma arising from mucosal sites in the head and neck (e.g., larynx, pharynx, oral cavity)	49	83
Malignant fibrous histiocytoma arising in soft tissues below the subcutaneous fat	5	8
Cancer of the lip	1	2
Cancer of the lung	1	2
Cancer of the kidney	1	2
Not cancer	2	3

TABLE 3—Cause of Death Listed on the Death Certificate and Correctness of Classification as Determined by Review of Records: Rhode Island Residents, 1979–1987

Diagnosis Listed on the Death Certificate	Misclassified Deaths	Correctly Classified Deaths
Squamous-cell (or epidermoid) cancer (or carcinoma):		
Of the neck or head and neck	33	3
Of other specific sites[a]	2	24
Not further specified	2	0
Cancer or carcinoma:		
Of the neck or head and neck	12	0
Of other specific sites[a]	1	7
Basal cell carcinoma	2	6
Malignant fibrous histiocytoma	6[b]	2
Other	1	9
Total	59	51

[a] Includes specific sites in the head and neck region (e.g., ear).
[b] Includes five deaths from malignant fibrous histiocytoma apparently arising deep to fascia.

The applicability of our findings to other areas of the United States is supported by the response of the Rhode Island Department of Health to our questionnaire. Their reported coding of cancer or squamous-cell cancer of the neck or of the head and neck was in accord with the practice reported by a majority of the health departments. A more thorough evaluation of the generalizability of our results awaits replication of our methods in other populations.

Prior investigations of the accuracy of NMSC death statistics are few. In 1970 through 1971, routinely coded death certificate diagnoses were compared with hospital diagnoses derived from the Third National Cancer Survey (TNCS).[8] Among the 67 deaths with the underlying cause NMSC (ICD-8 173) on the death certificate and with a hospital record of malignancy in the TNCS, only 18 (27%) were confirmed to have NMSC by the TNCS record abstract. These 18 also represented only 35% of the 51 deaths with any malignancy as the underlying cause and only NMSC noted in the hospital record abstract. Further evaluation of the medical records pertaining to these deaths was not performed.

Among 130 deaths coded as NMSC in California in 1959, 39 (30%) were determined to be due to NMSC, 73 (56%) were determined to be due to other causes, and, for 16 (12%), the cause could not be determined.[7] Among 133 NMSC deaths (i.e., coded under ICD-6 rubric 191) in Queensland, Australia, in 1960 through 1964, 95 (71%) on investigation were NMSC and 28 (21%) were other cancers.[9]

Most of the inaccuracies could be avoided by more consistent application of existing coding practices, by more uniform acceptance of the NCHS coding recommendations, and by changing the coding policy for squamous-cell cancer of the neck or head and neck to a rubric other than 173.

We must underscore the previously stated limitations in the interpretability of published NMSC mortality rates.[1,6,10,11]

TABLE 4—Coding Practices among the 41 Responding Health Departments[a] (United States, 1990)

Underlying Cause on Death Certificate:	ICD-9 Code	No. (%):
Malignant fibrous histiocytoma	173[b]	9 (22)
	171[c]	28 (68)
	216	4 (10)
Squamous-cell cancer of the neck	173[b,c]	38 (93)
	195	2 (5)
	199	1 (2)
Squamous-cell cancer of the head and neck (same responses as squamous-cell cancer of the neck)		
Cancer of the neck	173	2 (5)
	195[b,c]	39 (95)
Cancer of the head and neck (same responses as cancer of the neck)		

Note. ICD-9 = *International Classification of Diseases*, 9th version.
[a] In response to the question "Which ICD-9 code would you use to classify a death certificate with the following underlying cause listed on the death certificate?"
[b] Response of the Rhode Island Department of Health.
[c] The National Center for Health Statistics interpretation of the ICD-9 classification (M. Burt, personal communication, October 1990).

To best track the present epidemic of this very common cancer site, we need to complement existing sources of incidence data with refinements in the use of our primary source of mortality data: the death certificate. □

Acknowledgments

This study was supported by Department of Veterans Affairs medical research funds and National Cancer Institute Grant 49531.

A preliminary version was presented at the meeting of the Eastern Regional Society for Investigative Dermatology, New York City, October 1990.

We thank Margaret M. Boyle for her herculean efforts at locating medical records, and Roberta Chevoya, Joe Aldorissio, and John Fulton of the Rhode Island Department of Health for their assistance.

References

1. Boring CC, Squires TS, Tong T. Cancer statistics 1991. *CA* 1991;41:19–36.
2. Gallagher RP, Ma BM, McLean DI, et al. Trends in basal cell carcinoma, squamous cell carcinoma, and melanoma of the skin from 1973 through 1987. *J Am Acad Dermatol.* 1990;23:413–421.
3. Glass AG, Hoover RN. The emerging epidemic of melanoma and squamous cell skin cancer. *JAMA.* 1989;262:2097–2100.
4. Macdonald EJ, Heinze EB. *Epidemiology of Cancer in Texas: Incidence Analyzed by Type, Ethnic Group, and Geographic Location.* New York, NY: Raven Press; 1978.
5. Scotto J, Fears TR, Fraumeni JF. *Incidence of Nonmelanoma Skin Cancer in the United States.* Washington, DC: Public Health Service; 1983. NIH Publication No. 83–2433.
6. Weinstock MA. The epidemic of squamous cell carcinoma. *JAMA.* 1989;262: 2138–2140.
7. Dunn JE, Levin EA, Linden G, Harzfeld L. Skin cancer as a cause of death. *California Med.* 1965;102:361–363.
8. Percy C, Stanek E, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. *Am J Public Health.* 1981;71:242–250.
9. Gordon D, Silverstone H. Deaths from skin cancer in Queensland, Australia. In: Urbach F, ed. *The Biologic Effects of Ultraviolet Radiation.* New York: Pergamon Press; 1969:625–634.
10. Elwood JM, Lee JAH, Walter SD, Mo T, Green AES. Relationship of melanoma and other skin cancer mortality to latitude and ultraviolet radiation in the United States and Canada. *Int J Epidemiol.* 1974;3:325–332.
11. Lee JAH. The trend of mortality from primary malignant tumors of skin. *J Invest Dermatol.* 1973;59:445–448.

APPENDIX A—Criteria for Determining the Correctness of the Coded Cause of Death

Many of the records were traced and abstracted by a team of students (M.A., V.L., E.B., S.K.) in the last half of their final year of medical school. The remaining records were abstracted and the students' abstracts reviewed by a board-certified dermatologist (M.A.W.), who also adjudicated all ambiguities encountered by the medical student team and who was responsible for the final judgments regarding the correctness of the ICD classification according to the previously mentioned criteria. To resolve ambiguities in pathology reports, histologic specimens from certain biopsies were reviewed by a board-certified dermatopathologist (H.A.B.).

To accept a certificate as correctly coded under the ICD-9 rubric 173, we required that the diagnosis of nonmelanoma skin cancer be confirmed by medical records. We also required evidence from the record that the nonmelanoma skin cancer contributed significantly to the death. Some individuals did not die in medically supervised settings, or records of the events immediately leading to death were unavailable, so evidence of unresectable metastatic or advanced local disease (such as invasion of the orbital bones) was accepted as indicative of a significant contribution of the tumor to death. If the patient's physician believed that resection of extensive tumor was not feasible or not indicated because of the patient's poor prognosis, we considered the disease unresectable. We did not require that the available documentation of symptoms, signs, and tests convince us that the nonmelanoma skin cancer was the most important underlying cause of death. This degree of detail was often absent from the available documentation.

We determined that a certificate was incorrectly coded under the ICD-9 rubric 173 if we concluded from the record that the malignancy responsible for death was melanoma or arose from a noncutaneous primary site, or if the nonmelanoma skin cancer appeared unrelated to the events leading to death.

APPENDIX B—Description of the Six Deaths Excluded from Analysis

Two of the deaths involved individuals with aggressive AIDS-associated Kaposi's sarcoma. Since the ICD classification for this cause had changed from 173 to 042, we excluded these from the analysis. We were unable to verify or disprove the coded cause of death on the basis of medical records for four deaths. Complete records were obtained for two of these cases, but the primary site of the malignancy that led to death was unclear. In another two cases, we were unable to obtain complete medical records, either because the records were lost or destroyed or because the source of medical care during that time could not be traced. The nursing home at which one of these two patients last resided refused access to her medical records.

Provider Compliance with Recommended Dietary Fluoride Supplement Protocol

ABSTRACT

Proper dosage of dietary fluoride supplements is important for children without access to optimally fluoridated water to balance dental caries prevention and risks of dental fluorosis. Subsequent to water fluoride assay, compliance by physicians and dentists with the recommended dietary fluoride supplement dosage schedule was determined for 446 children. Approximately one third of child patients and 42% of siblings did not receive the recommended supplement dosage. (*Am J Public Health.* 1992;82:281–283)

Steven M. Levy, DDS, MPH, and Gary Muchow, MA, PhD

Introduction

Community water fluoridation remains the most efficient and cost-effective method to receive the recommended levels of fluoride for dental caries prevention.[1] Dietary fluoride supplements are advocated[2] for children not served by optimally fluoridated water (46% in the United States).[3]

Before prescribing supplements, fluoride histories and assays of the major sources of drinking water should be obtained in order to establish a balance between dental caries prevention and increased risks of dental fluorosis (mottled enamel).[2] Dental fluorosis is an enamel hypoplasia caused by the ingestion of excessive fluoride during the years of tooth calcification.[4] Water fluoride assay as a tool in preventing dental fluorosis is especially important in light of recent evidence of an increased prevalence of dental fluorosis,[4–6] of substantial prevalence nationwide,[7] and of the association of dietary fluoride supplement use with increased fluorosis risks.[6,8–11] The need for water fluoride assay to help avoid fluorosis has been emphasized recently in the US Department of Health and Human Services' review of the benefits and risks of fluoride.[12]

Table 1 shows the dietary fluoride supplement dosage schedule currently recommended.[13,14] Dosage is dependent on age and water fluoride level. There are many steps in the proper evaluation of dietary fluoride supplement need, dosage determination, and monitoring of patient compliance.[15] Many studies have reported difficulties with patient compliance,[16] but few have addressed provider compliance with recommended protocol.[17,18] A pilot study investigating the interactions that occurred between provid-

The authors are with the Department of Preventive and Community Dentistry, College of Dentistry, University of Iowa, Iowa City.

Requests for reprints should be sent to Steven M. Levy, DDS, MPH, Associate Professor, Department of Preventive and Community Dentistry, College of Dentistry, University of Iowa, Iowa City, IA 52242.

This paper was submitted to the journal October 15, 1990, and accepted with revisions June 10, 1991.

TABLE 1—Recommended Supplemental Fluoride Dosage Schedule[a] (in mg of fluoride per day)

Age of Child, y	Parts per Million of Fluoride in Water Supply		
	<0.3	0.3 to 0.7	>0.7
Birth to 2[b]	0.25	0	0
2 to 3	0.50	0.25	0
3 to 13[b]	1.00	0.50	0

[a]Recommended by the Council on Dental Therapeutics of the American Dental Association and the Committee on Nutrition of the American Academy of Pediatrics.
[b]The American Academy of Pediatrics recommends providing supplementation from 2 weeks of age through at least 16 years of age.

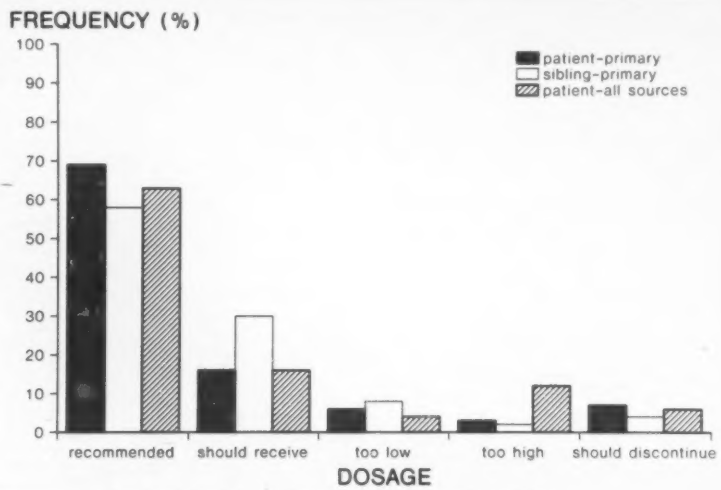


FIGURE 1—Appropriateness of dietary fluoride supplementation.

ers and patients subsequent to the submission of water samples for fluoride assay indicated that a substantial proportion of children were not receiving the appropriate levels of fluoride supplements.[15] The purpose of this study was to investigate further these patterns.

Methods

The State of North Carolina Laboratory of Public Health, the University of Iowa College of Dentistry, and the Iowa (State) Hygienic Laboratory provide assays of water samples for fluoride content, with results reported to the nearest 0.01 parts per million (ppm) of fluoride. The North Carolina laboratory uses an automated complexone method with distillation,[19] whereas the other two laboratories use fluoride ion-specific electrode techniques. The state laboratories are approved by the Environmental Protection Agency, and results have consistently been within 0.05 ppm of readings of Centers for Disease Control (CDC) blind standards.[20] The University of Iowa College of Dentistry does not participate in the CDC program but regularly performs several internal quality assurance checks. These three laboratories were chosen because of the differences in the individuals to whom results are reported. (The North Carolina laboratory reports results only to health care providers, Iowa College of Dentistry results may go to provider or patient, and Iowa state laboratory results usually are sent only to the patient, since the assay request form does not include provider information.)

Results of 647 samples submitted for fluoride assay between March 1987 and February 1989 were analyzed. Parents of 548 patients (or 85%) were successfully contacted 3 to 4 months after the assay results were sent, permitting provider/patient follow-up to occur before our survey was administered. Thirty-nine percent of the samples were analyzed at the North Carolina state laboratory, 36% at the Iowa College of Dentistry, and 24% at the Iowa Hygienic Laboratory.

Results

Subsequent to the reporting of the assay results, contacts occurred between parents and health care providers in 85% of the cases. Eight percent were with a second provider, either instead of or in addition to the provider who recommended the assay (e.g., a physician recommended assay and a dentist was contacted to assist with interpretation of the results).

More than half of the parents reported that their child received substantial amounts of water outside of the home (34% child care and 22% school), but only 29% reported that the provider inquired about additional sources of water.

Prescriptions for dietary fluoride supplements (tablets or drops) were provided for 75% of the patients. Eighteen percent were prescribed prior to the assay, with the provider probably assuming negligible levels of fluoride in the drinking water and prescribing full dosage supplements. However, only about half of the samples in Iowa and 80% to 85% in North Carolina had fluoride content less than 0.3 ppm.[15,21,22]

Using patient age, assay results, and recommended dosage schedules, determinations were made as to the appropriateness of the fluoride supplement received by patients and siblings based on their reported primary source of water. Because of incomplete information, such determinations were not possible for 19% of the patients and 12% of the siblings. The results (Figure 1) exclude these "unknowns."

Sixty-nine percent of the patients received the recommended dosage (including no supplement, when indicated) after the assay results had been reported. Sixteen percent were not receiving a supplement when indicated, and 6% needed a supplement at a higher dosage. Three percent should have received reduced dosage, and 7% needed supplementation discontinued.

Fifty-eight percent of the siblings received a correct dosage. Thirty percent of siblings were not receiving a fluoride supplement when one was indicated. These siblings could have received appropriate supplementation without further assay of their home drinking water being necessary. Eight percent needed a supplement at a higher dosage, 2% needed a lower dosage, and 4% needed supplementation discontinued.

Appropriateness of supplementation was also determined for patients based on weighting of multiple patient sources of drinking water (e.g., home, school, and day care). This determination of appropri-

ateness was made only in those cases where additional fluoride assay was not necessary because the additional sources involved community water systems of known fluoride content. Sufficient information to determine appropriateness in this way was available for 342 patients. Among these, 63% received a correct dosage of fluoride. Sixteen percent did not receive a supplement when one was indicated, 4% needed a higher dosage, 12% needed a lower dosage, and 6% needed to stop supplementation.

Discussion

The findings are consistent with those of our previous pilot study. Approximately one third of patients for whom water fluoride assay was performed did not receive the recommended dosage of the dietary fluoride supplement. Among these, about one third were at increased risk of dental fluorosis and two thirds were not receiving the recommended level for caries prevention. Siblings were less likely to be receiving the recommended supplement dosage. These results likely are conservative estimates of discrepancies among all providers, because the patients and providers in this study probably were more knowledgeable and/or motivated, as evidenced by their assay efforts.

Since more children are receiving substantial proportions of drinking water outside the home, patient water sources other than at home were considered. This often was important when home drinking water was nonfluoridated but water at child care or school was optimally fluoridated. Following the "standard practice" of considering fluoride from the home source only and providing a dietary fluoride supplement would place the child at higher risk of dental fluorosis. Thus, as suggested previously,[21,22] it probably is now appropriate to consider all important water sources, including those other than home, in the standard approach and to weight proportionately the ingestion from multiple water sources and the fluoride levels. As always, adjustments of supplement dosage are necessary as children turn ages 2 and 3, even where fluoride content of important water sources remains the same.

The reporting systems in this study supplied the provider or parent with water fluoride assay results but not with information for interpretation of the results. Most laboratory reports for other types of analyses provide information about the range of "normal" laboratory values. In order for appropriate fluoride supplements to be prescribed more consistently, it may be advisable for laboratories to include with their assay results an insert with dosage schedule and guidelines for interpretation of results. In addition, public health personnel must consider the cost–benefit ratio of dietary fluoride supplement and water assay programs. Especially in those states without existing water fluoride assay programs,[23] it may be appropriate for dietary fluoride supplements and water fluoride assay to be used only with children at high risk of dental caries. □

Acknowledgments

This project was supported by US Public Health Service Grant 2 S07-RR5313-24, the American Fund for Dental Health, the North Carolina Division of Dental Health, and the University of Iowa Spelman Rockefeller Seed Grant Program for Children.

Portions of this paper were presented at the annual meetings of the American Association of Public Health Dentistry (1989) and the North American Primary Care Research Group (1990).

References

1. Burt BA, ed. Cost-effectiveness of caries prevention in dental public health. Results of the workshop. *J Public Health Dent.* 1989;49:331–340.
2. Levy SM. Expansion of the proper use of systemic fluoride supplements. *J Am Dent Assoc.* 1986;112:30–34.
3. Centers for Disease Control. *Fluoridation Census 1985.* Atlanta, Ga: Centers for Disease Control; 1988.
4. Szpunar SM, Burt BA. Trends in the prevalence of dental fluorosis in the United States: a review. *J Public Health Dent.* 1987;47:71–79.
5. Heifetz SB, Driscoll WS, Horowitz HS, Kingman A. Prevalence of dental caries and dental fluorosis in areas with optimal and above-optimal water-fluoride concentrations. A 5-year follow-up survey. *J Am Dent Assoc.* 1988;116:490–495.
6. Kumar J, Green E, Wallace W, Carnahan T. Trends in dental fluorosis and dental caries prevalence in Newburgh and Kingston, NY. *Am J Public Health.* 1989;79:565–570.
7. Brunelle JA. The prevalence of dental fluorosis in U.S. children, 1987. *J Dent Res.* 1989;68(special issue):995. Abstract.
8. Leverett DH, Adair SM, Proskin HM. Dental fluorosis among children in fluoridated and non-fluoridated communities. *J Dent Res.* 1988;67(special issue):230. Abstract.
9. Woolfolk MW, Faja BW, Bagramian RA. Relation of sources of systemic fluoride to prevalence of dental fluorosis. *J Public Health Dent.* 1989;49:78–82.
10. Bohaty DG, Parker WA, Seale NS, Zimmerman ER. The prevalence of fluorosis-like lesions associated with topical and systemic fluoride usage in an area of optimal water fluoridation. *Pediatr Dent.* 1989;11:125–128.
11. Pendrys DG, Katz RV. Risk of enamel fluorosis associated with fluoride supplementation, infant formula, and fluoride dentifrice use. *Am J Epidemiol.* 1989;130:1199–1208.
12. Subcommittee on Fluoride of the Committee to Coordinate Environmental Health and Related Programs. *Review of Fluoride—Benefits and Risks.* Washington, DC: Public Health Service; US Department of Health and Human Services; 1991.
13. American Academy of Pediatrics, Council on Nutrition. Fluoride supplementation. *Pediatrics.* 1986;77:758–761.
14. American Dental Association, Council on Dental Therapeutics. Fluoride Compounds. In: *Accepted Dental Therapeutics.* 40th ed. Chicago, Ill: American Dental Association; 1984.
15. Levy SM, Carrell AF. Compliance by health care providers with recommended systemic fluoride supplementation protocol. *Clin Prev Dent.* 1987;9:19–22.
16. Newbrun E. Water fluoridation and fluoridation supplements in caries prevention. *J Calif Dent Assoc.* 1980;8:38–47.
17. Kuthy RA, McTigue DJ. Fluoride prescription practices of Ohio physicians. *J Public Health Dent.* 1987;47:172–176.
18. Levy SM, Rozier RG, Bawden JW. Use of systemic fluoride supplements by North Carolina dentists. *J Am Dent Assoc.* 1987;114:347–350.
19. Clesceri LS, Greenberg AE, Trussell RR, eds. *Standard Methods for the Examination of Water and Wastewater.* 17th ed. Washington, DC: American Public Health Association; American Water Works Association; Water Pollution Control Federation; 1989.
20. Centers for Disease Control. *CDC: Fluoridated Drinking Water, Proficiency Testing.* Atlanta, Ga: Centers for Disease Control, 1988.
21. Levy SM, Cipriano KB, Maurer WC. Systemic fluoride supplements for Iowa children—the dentists' role. *Iowa Dent J.* 1985;71:33–37.
22. Levy SM, Bawden JW, Bowden BS, Rozier RG. Fluoride analyses of patient water supplies requested by North Carolina health professionals. *Am J Public Health.* 1984;74:1412–1413.
23. Levy SM, Shavlik DA. The status of water fluoride assay programs and implications for prescribing of dietary fluoride supplements. *J Dent Child.* 1991;58:23–26.

A2623-1

American Journal of Public Health

March 1992, Vol. 82, No. 3 Established 1911

86389
028
CONGRESSIONAL INFORMATION SVC
4520 EAST-WEST HIGHWAY
BETHESDA MD 20814
ED KIANKA

APHA
American Public Health Association

Featuring Children's and Adolescents' Health

Editorials

341 Wanting Children *C. A. Miller*

343 Preterm Birth, Intrauterine Growth Retardation, and Perinatal Mortality *J. L. Kiely and M. Susser*

Public Health Policy Forum

346 Nutrition: The Need to Define "Optimal" Intake as a Basis for Public Policy Decisions *E. L. Wynder, J. H. Weisburger, and S. K. Ng*

Commentary

351 Snuffing Tobacco Out of Sport *G. N. Connolly, C. T. Orleans, and A. Blum*

Featuring Children's and Adolescents' Health

354 Prevalence of Obesity among Children of Military Dependents at Two Major Medical Centers *C. M. Tiwary and A. H. Holguin*

358 Body Fatness and Risk for Elevated Blood Pressure, Total Cholesterol, and Serum Lipoprotein Ratios in Children and Adolescents *D. P. Williams, S. B. Going, T. G. Lohman, D. W. Harsha, S. R. Srinivasan, L. S. Webber, and G. S. Berenson*

364 Childhood Chronic Illness: Prevalence, Severity, and Impact *P. W. Newacheck and W. R. Taylor*

372 Reductions in HIV Risk–Associated Sexual Behaviors among Black Male Adolescents: Effects of an AIDS Prevention Intervention *J. B. Jemmott III, L. S. Jemmott, and G. T. Fong*

378 Birth Weight and Perinatal Mortality: The Effect of Gestational Age *A. J. Wilcox and R. Skjærven*

422 Morbidity among Pediatric Motor Vehicle Crash Victims: The Effectiveness of Seat Belts *J. S. Osberg and C. Di Scala*

426 Pediatric Gonococcal Infection, Florida, 1984 to 1988 *J.-C. A. Desenclos, D. Garrity, and J. Wroten*

429 Evaluation of an Intervention to Reduce Playground Hazards in Atlanta Child-Care Centers *J. J. Sacks, M. D. Brantley, P. Holmgreen, and R. W. Rochat*

432 Air Quality during the Winter in Québec Day-Care Centers *S. Daneault, M. Beausoleil, and K. Messing*

435 Infant Mortality, Per Capita Income, and Adult Illiteracy: An Ecological Approach *R. Tresserras, J. Canela, J. Alvarez, J. Sentis, and L. Salleras*

438 Which Newborns in New York City Are at Risk for Special Education Placement? *D. Goldberg, M. McLaughlin, M. Grossi, A. Tytun, and S. Blum*

441 The Lowest Birth-Weight Infants and the US Infant Mortality Rate: NCHS 1983 Linked Birth/Infant Death Data *M. D. Overpeck, H. J. Hoffman, and K. Prager*

445 An Epidemiological Comparison of Mexican-American and White Non-Hispanic 8th- and 12th-Grade Students' Substance Use *E. L. Chavez and R. C. Swaim*

(continued page 331)

Contents

American Journal of Public Health

"Information for Authors" appeared on page 138 of the January 1992 issue and may also be obtained from the Journal Editorial Office. An abbreviated version is printed in this issue.

American Journal of Public Health
1015 Fifteenth Street, NW
Washington, DC 20005

APHA
American Public Health Association

March 1992, Vol. 82, No. 3 ISSN: 0090-0036 Established 1911

Articles

383 Estimation of the Break-Even Point for Smoking Cessation Programs in Pregnancy *M. Shipp, M. S. Croughan-Minihane, D. B. Petitti, and A. E. Washington*

391 Low Birth Weight in Relation to Multiple Induced Abortions *M. T. Mandelson, C. B. Maden, and J. R. Daling*

395 Mortality, Disability, and Falls in Older Persons: The Role of Underlying Disease and Disability *J. E. Dunn, M. A. Rudberg, S. E. Furner, and C. K. Cassel*

401 Living Arrangements and Survival among Middle-Aged and Older Adults in the NHANES I Epidemiologic Follow-up Study *M. A. Davis, J. M. Neuhaus, D. J. Moritz, and M. R. Segal*

407 Readmission after Surgery in Washington State Rural Hospitals *H. G. Welch, E. H. Larson, L. G. Hart, and R. A. Rosenblatt*

412 The Behavioral Risk Factor Survey and the Stanford Five-City Project Survey: A Comparison of Cardiovascular Risk Behavior Estimates *C. Jackson, D. E. Jatulis, and S. P. Fortmann*

417 Smokeless Tobacco, Cardiovascular Risk Factors, and Nicotine and Cotinine Levels in Professional Baseball Players *D. Siegel, N. Benowitz, V. L. Ernster, D. G. Grady, and W. W. Hauck*

Public Health Briefs

448 Analyzing the Effect of Prenatal Care on Pregnancy Outcome: A Conditional Approach *M. H. Malloy, T.-C. Kao, and Y. J. Lee*

450 The Safety of Home Birth: The Farm Study *A. M. Durand*

453 Skilled Nursing Facility Care for Persons with AIDS: Comparison with Other Patients *J. H. Swan, A. E. Benjamin, and A. Brown*

455 Risk Factors Associated with the Classification of Unspecified and/or Unexplained Causes of Death in an Occupational Cohort *D. L. Cragle and A. Fetcher*

458 Performance of the Reflotron in Massachusetts' Model System for Blood Cholesterol Screening Program *S. Havas, R. Bishop, L. Koumjian, J. Reisman, and S. Wozenski*

Other Journal Departments

332 Masthead
462 Information for Authors
336 Advertisers' Index
463 Letters to the Editor
468 Book Corner
472 Reviewer List
476 Association News: APHA Policy Statements
495 Membership Application
496 Job Opportunities

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.

APHA OFFICERS

Joyce C. Lashof, MD, *President*
C. William Keck, MD, MPH, *Immediate Past President*
Helen Rodriguez-Trias, MD, *President-Elect*
Arnold E. Greenberg, SM, *Vice-President (USA)*
Trevor Hancock, MB, MHSc, *Vice-President (Canada)*
Julio Teja Perez, MD, *Vice-President (Latin America)*
Caswell A. Evans Jr, DDS, MPH, *Treasurer*
Frances M. Veverka, RS, MPH, *Chair, Executive Board*
George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*
William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD

Frances M. Veverka, RS, MPH (1992), *Chair*
Marilyn Aguirre-Molina, EdD, MS (1993)
Brian A. Dobrow, MPH (1992)
Caswell A. Evans, Jr, DDS, MPH, *Treasurer*
Kristine M. Gebbie, BSN, MN (1994), *Vice-Chair*
Alan R. Hinman, MD, MPH (1995)
Janet Heinrich, DrPH, MPH (1993)
Margaret H. Jordan, RN, MPH (1992)
C. William Keck, MD, MPH, *Immediate Past President*
Margo D. Kerrigan, MPH (1995)
Joyce C. Lashof, MD, *President*
Barry S. Levy, MD, MPH (1993)
Helen Rodriguez-Trias, MD, *President-Elect*
Allan G. Rosenfield, MD (1994)
Reed V. Tuckson, MD (1994)
Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO

William H. McBeath, MD, MPH, *Executive Director*
Joanne E. Lukomnik, MD, *Chair, Action Board*
George Pickett, MD, MPH, *Chair, Program Development Board*
George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc., 1015 Fifteenth St., NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50,000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health,* 1015 Fifteenth St., NW, Washington, DC 20005.

Information on Manuscript Preparation is published in this Journal in January. The Journal publishes reports of original research, demonstrations, evaluations, and other articles covering current aspects of public health. Manuscripts are accepted for consideration with the understanding that they, or their essential substance, have not been published previously or accepted for publication and are not under consideration elsewhere.

Subscription price, individual, is $80/year US, $120/year international; **institutional,** $160/year US, $240/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to American Journal of Public Health, 1015 Fifteenth St., NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact: Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact: Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols. 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS Bibliographic Retrieval Service Inc, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc., Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Body Fatness and Risk for Elevated Blood Pressure, Total Cholesterol, and Serum Lipoprotein Ratios in Children and Adolescents

Daniel P. Williams, MS, Scott B. Going, PhD, Timothy G. Lohman, PhD, David W. Harsha, PhD, Sathanur R. Srinivasan, PhD, Larry S. Webber, PhD, and Gerald S. Berenson, MD

ABSTRACT

Background. Recent studies have shown considerable variation in body fatness among children and adolescents defined as obese by a percentile rank for skinfold thickness.

Methods. We examined the relationship between percent body fat and risk for elevated blood pressure, serum total cholesterol, and serum lipoprotein ratios in a biracial sample of 3320 children and adolescents aged 5 to 18 years. Equations developed specifically for children using the sum of subscapular (S) and triceps (T) skinfolds were used to estimate percent fat. The S/T ratio provided an index of trunkal fat patterning.

Results. Significant overrepresentation (>20%) of the uppermost quintile (UQ) for cardiovascular disease (CVD) risk factors was evident at or above 25% fat in males (32.2% to 37.3% in UQ) and at or above 30% fat in females (26.6% to 45.4% in UQ), even after adjusting for age, race, fasting status, and trunkal fat patterning.

Conclusions. These data support the concept of body fatness standards in White and Black children and adolescents as significant predictors of CVD risk factors. Potential applications of these obesity standards include epidemiologic surveys, pediatric health screenings, and youth fitness tests. (*Am J Public Health.* 1992;82:358–363)

Introduction

The association between skin-fold thicknesses in children and adolescents and elevated levels of variables considered to be indicative of risk for cardiovascular disease (CVD) in adults has been found by several investigators.[1–5] In population-based studies of children and adolescents, for example, excess subcutaneous fatness has been associated with elevated blood pressure (BP),[1,2,6] serum lipids, and lipoprotein fractions,[2–4] and the clustering of these variables.[5] Although the results of these studies demonstrate that elevated levels of CVD risk factors often accompany excess body fat in youth, the absolute level of body fat corresponding to significant risk for high BP, serum lipids, low-density lipoprotein (LDL) and very low-density lipoprotein (VLDL) cholesterols is not known.

Although children and youth with triceps skinfolds ≥ the 85th percentile for age and sex are at greater risk for high BP than children with lower skinfold thicknesses,[1] such a rank-ordered definition of obesity implies that the absolute level of body fatness associated with increased risk is variable. Recent work has shown that the 85th percentile may not be indicative of obesity in all samples, because skinfold thicknesses at a given percentile correspond to different levels of body fat at different ages.[7,8] The 85th skinfold percentile, for example, represents fatness levels ranging from 17% to 22% in children and from 25% to 34% in adolescents.[8]

In addition to the discrepancy between relative and absolute definitions of obesity, the triceps skinfold site has been shown to be unrelated to BP in youth when statistical adjustments were made for subscapular skinfold thickness.[9] Moreover, a trunkal fat pattern, independent of general body fatness, is related to high levels of LDL and VLDL and to low levels of high-density lipoprotein (HDL) cholesterol in 6- to 18-year-old youth.[10] The independent effects of centripetal localization of subcutaneous fat on BP[9,11] and adverse lipoprotein profiles[10] make it important to control the potential confounding effect of fat patterning when the relation between total body fat and CVD risk factors in youth is evaluated.

The purpose of the present study was to determine the level of body fatness, independent of fat patterning, associated with significant risk for elevated BP, total cholesterol, and serum lipoprotein ratios in a biracial sample of children and adolescents aged 5 to 18 years. The aim was to develop cardiovascular health-related percent body fat standards that may be applied to epidemiologic investigations of the prevalence and incidence of obesity in children and adolescents, pediatric health screenings, and national youth fitness tests.[12,13]

Daniel P. Williams, Scott B. Going, and Timothy G. Lohman are with the Department of Exercise and Sport Sciences, University of Arizona, Tucson. David W. Harsha, Sathanur R. Srinivasan, Larry S. Webber, and Gerald S. Berenson are with the Department of Medicine, Louisiana State University Medical Center, New Orleans.

Requests for reprints should be sent to Daniel P. Williams, MS, Department of Exercise and Sport Sciences, 101 Gittings Bldg, University of Arizona, Tucson, AZ 85721.

This paper was submitted to the Journal March 28, 1991, and accepted with revisions September 13, 1991.

Methods

Study Sample

The sample included 1667 males (1062 White, 605 Black) and 1653 females (1028 White, 625 Black) aged 5 to 18 years. All subjects were residents of Ward 4, Washington Parish, La, and were participants in the Bogalusa Heart Study. The data were collected in the years 1983 through 1985. An extensive review of the demographic characteristics of this biracial community, as well as the overall design of this large, population-based investigation of children and youth, has been published elsewhere.[14] The Louisiana State University Medical Center reviews experiments related to human subjects. Written informed consent was obtained from a parent or guardian of each child.

Procedures

Total body fat and fat distribution. Total body and regional fatness were estimated from limb (triceps) and trunk (subscapular) skinfold thicknesses. Lange calipers were used to measure triceps and subscapular skinfolds to the nearest 1.0 mm. Skinfold measurement technique and site location have been described previously.[14,15] The reliability of these measurements, assessed by test-retest intraclass correlations, was of equal magnitude ($r = 0.98$) for both sites.[16]

Body density (D_b) was estimated from age and the sum of triceps and subscapular skinfolds (STSS) and was subsequently used to derive total percent body fat. The STSS-to-D_b conversion was based on race- and gender-specific regression equations that were derived from the pooled skinfold and body density data from several samples of White and Black youth measured in different laboratories (Appendix).[17–20] We then used a modification of the density–to–percent-fat conversion constants reported by Lohman[21] to convert D_b to percent body fat (Table 1, Eqs. 1 and 2). These youth-specific constants adjust for growth- and maturation-related differences in fat-free body density (D_{FFB}) at different ages and prevent the systematic overestimation of body fat when adult density–to–percent-fat equations are applied to chemically immature children and youth.[22] The present combination of equations was chosen over other multicomponent-based equations for youth that convert STSS directly to percent fat[23] because the multilaboratory equations used herein may be more generalizable.

TABLE 1—Multilaboratory-based Estimation of Body Density from the Sum of Triceps and Subscapular Skinfolds (STSS) and Age in White and Black Children and Adolescents[a]

Variable	Males		Females	
	White (n = 262)	Black (n = 96)	White (n = 219)	Black (n = 94)
STSS (mm)	−0.00227	−0.00203	−0.00188	−0.00162
STSS^2 (mm^2)	0.000015	0.000012	0.000013	0.000008
Age, yr	0.00243	0.00193	0.00191	0.00165
Intercept	1.0600	1.0682	1.0533	1.0580
Adjusted intercept		1.0666		1.0555
R^2	0.70	0.65	0.58	0.70
SEE (g/cm^3)	0.0102	0.0107	0.0096	0.0089

[a]Because currently available equations for converting body density (D_b) to percent fat[21] are based more on the fat-free body composition of Whites than of Blacks, we adjusted the intercepts in the above equations for Black males and females to remove the influence of differences in fat-free body density (D_{FFB}) on D_b and, subsequently, on estimates of body fat. Body density was converted to body fat using age-adjusted, gender-specific equations derived by regression analysis from age and the recently reported age-specific body density constants,[21] thereby deriving the following equations to estimate percent fat (PF) from D_b:

(1) $PF_{males} = [(5.68 - (0.041 \times age)) \div D_b] - [(5.31 - (0.045 \times age))] \times 100$

(2) $PF_{females} = [(5.69 - (0.038 \times age)) \div D_b] - [(5.31 - (0.041 \times age))] \times 100$

To determine the critical level of body fat associated with elevated CVD risk factor variables, the males and females were subsequently grouped by level of percent fat. Males were divided into the following five fatness groups: <10% (n = 217), 10% to 14.9% (n = 575), 15% to 19.9% (n = 436), 20% to 24.9% (n = 192), and ≥25% (n = 247). Females were grouped by body fat as follows: <20% (n = 547), 20% to 24.9% (n = 493), 25% to 29.9% (n = 313), 30% to 34.9% (n = 192), and ≥35% (n = 108).

An index of trunk-to-limb fat distribution was computed as the natural logarithm of the subscapular/triceps skinfold thickness ratio.[24] Because absolute standards for children and youth do not exist for this index of fat patterning, age-, sex- and race-specific percentiles for the present sample were developed. Fat-patterning quintile, used as a covariate of percent fat in this report, refers to the contrast of youth in the uppermost quintile for $\log_n$ (subscapular/triceps) with those in the lower four quintiles. This contrast was used in the analysis to control for the potential confounding effect of trunkal fatness in assessing the relation of total body fat with BP[9] and serum lipoproteins.[10]

CVD risk factor variables. The protocol, standardization, and training of observers used to measure BP[14] and the blood collections required to assess serum total cholesterol and lipoprotein levels[3] have previously been outlined in detail. Seated, resting BP measures on the right arm were made in triplicate with a mercury sphygmomanometer by trained nurses; the mean of six readings at the first (systolic) and fourth (diastolic) Korotkoff sounds was used in the data analysis. In an earlier Bogalusa Heart Study survey, reliability coefficients ranged from $r = .84$ to $r = .86$ for systolic BP and from $r = .76$ to $r = .78$ for diastolic BP.[16]

Serum total cholesterol and lipoproteins were measured at Louisiana State University Core Lipid Laboratory, New Orleans, La, which has been standardized by the Centers for Disease Control in Atlanta, Ga. Laboratory measurement quality is routinely monitored by a surveillance program. Serum total cholesterol concentration was measured with an AutoAnalyzer II, according to the recommendations of the Lipid Research Clinics.[25] A combination of heparin-calcium precipitation and agar-agarose gel electrophoresis[26] was used to determine the levels of lipoprotein cholesterol fractions (VLDL-C, LDL-C, and HDL-C). Reliability coefficients (intraclass correlations) ranging from $r = .89$ to $r = .97$ were reported earlier for these variables.[16]

For the purpose of this report, elevated BP, total cholesterol, and lipoprotein cholesterols were individually defined as levels in the uppermost quintile of the age-, race-, and sex-specific distributions for systolic BP, diastolic BP, total cholesterol, the LDL-C/HDL-C ratio, and the ratio of VLDL-C + LDL-C/HDL-C. The uppermost quintile was chosen to define elevated CVD risk factor variables because an earlier study[27] demonstrated this cutoff point to be a more sensitive indica-

TABLE 2—Sample Size and Percentage of Subjects in the Fifth Quintile (Q5) for Blood Pressure, Total Cholesterol, and Serum Lipoprotein Ratios at Increasing Levels of Body Fat

	SBP[a]			DBP[b]			TCHOL[c]			LR-1[d]			LR-2[e]		
% Fat	No. Q1–5	No. Q5	% Q5	No. Q1–5	No. Q5	% Q5	No. Q1–5	No. Q5	% Q5	No. Q1–5	No. Q5	% Q5	No. Q1–5	No. Q5	% Q5
Males															
<10.0	217	17	7.8*	217	26	12.0*	213	39	18.3	213	37	17.4	213	34	16.0
10.0–14.9	575	99	17.2*	573	87	15.2*	570	91	16.0*	566	89	15.7*	566	84	14.8*
15.0–19.9	436	85	19.5	436	82	18.8	432	83	19.2	432	80	18.5	432	78	18.1
20.0–24.9	192	40	20.8	192	47	24.5**	191	39	20.4	191	33	17.3	191	34	17.8
≥25.0	247	86	34.8**	247	86	34.8**	236	76	32.2**	228	82	36.0**	228	85	37.3**
Females															
<20.0	547	68	12.4*	547	66	12.1*	538	88	16.4*	537	77	14.3*	537	80	14.9*
20.0–24.9	493	87	17.6*	493	85	17.2*	481	78	16.2*	478	82	17.2*	478	83	17.4*
25.0–29.9	313	75	24.0	313	73	23.3	308	72	23.4	308	52	16.9	308	54	17.5
30.0–34.9	192	51	26.6**	192	61	31.8**	187	54	28.9**	187	61	32.6**	187	62	33.2**
≥35.0	108	49	45.4**	108	40	37.0**	100	31	31.0**	99	35	35.4**	99	37	37.4**

[a]SBP = systolic BP.
[b]DBP = diastolic BP.
[c]TCHOL = total cholesterol.
[d]LR-1 = LDL-C/HDL-C.
[e]LR-2 = (VLDL-C + LDL-C)/HDL-C.
*significantly less than that expected by chance, $P < .05$.
**significantly greater than that expected by chance, $P < .05$.

tor of future high BP than was the more stringent 95th percentile.

Data analysis. A chi-square test was used to determine whether the percentages of subjects in the fifth quintile for BP, total cholesterol, and lipoprotein ratios were significantly different than those expected by chance alone (20%) at each level of body fatness in both sexes. Logistic regression analysis[28] was used to estimate the relative odds for elevated BP, total cholesterol, and lipoprotein ratios between contrasting percent fat groups of interest, while holding potential confounders constant. Presence or absence of elevated BP, total cholesterol, and lipoprotein ratios were the dependent variables in the logistic regression analyses. The independent variables included several percent fat group contrasts; a contrast for fat-patterning quintile; race; age; and, for the total cholesterol and lipoprotein ratio analyses, fasting status. The risk for membership in the uppermost quintile for the cardiovascular variables was defined as 1.00 in the lowest percent fat group. The odds of uppermost quintile membership were thus determined for each of the upper four percent fat groups relative to the leanest group within each gender.

Results

The unadjusted bivariate relationships between body fat and CVD risk factor variables is shown in Table 1. A dose-response effect was observed between BP and percent body fat; in contrast, total cholesterol and the lipoprotein ratios were relatively independent of percent body fat among the lower four fatness groups in males and the lower three fatness groups in females (Table 2). The percentage of subjects in the uppermost quintile for all CVD risk factor variables was significantly ($P < .05$) greater than that expected by chance alone (20%) in males with ≥25% fat and in females with ≥30% fat. The ≥35% fat group of females displayed even greater overrepresentation (31.0% to 45.4% vs 26.6% to 33.2%) in the uppermost quintile than did the 30% to 34.9% fat group of females.

Multivariate-based estimates of the relative odds for elevated BP, total cholesterol, and lipoprotein ratios, adjusted for potential confounding variables at increasing levels of percent body fat, are presented in Figure 1 for males and Figure 2 for females. After accounting for potential confounding by age, race, trunkal fat patterning, and fasting status, males with body fat ≥25.0% were 2.8 (1.7–4.8) to 7.0 (3.6–13.6) times as likely as those in the leanest group (<10% fat) to have elevated systolic BP, diastolic BP, LDL-C/HDL-C ratios, and VLDL-C + LDL-C/HDL ratios (Figure 1). In females, the 30% to 34.9% fat group was 2.7 (1.6–4.5) to 3.8 (2.3–6.1) times more likely than the leanest group (<20% fat) to be in the uppermost quintile for these CVD risk factor variables; the likelihood for elevated BP and lipoprotein ratios in the female ≥35.0% fat group ranged from 3.3 (1.8–5.9) to 4.9 (2.8–8.7) (Figure 2).

Discussion

This study is unique because critical levels of body fat, as opposed to skinfold thickness percentile, that correspond to significant risk for elevated BP and lipoprotein ratios were determined and found to be independent of trunkal fat patterning in a biracial sample of children and youth. Inherent in previous normative value-derived definitions of childhood obesity[1] is the considerable age-related variation in body fatness at a constant percentile for skinfold thickness.[7,8] This concept is supported by the finding that the 85th percentile for the STSS in the present sample of children and adolescents represents body fatness values ranging from 14.8% to 34.3% fat in males and from 21.9% to 37.7% fat in females. Moreover, the finding that critical levels of body fatness were related to elevated levels of CVD risk factor variables independent of trunkal fat patterning is particularly relevant in light of recent studies[5,10,29] establishing the link between trunk fatness and high levels of BP, serum lipoprotein fractions, and serum insulin in children and adolescents. The independent relationships of CVD risk factors with percent body fat in children and adolescents and with trunkal fat patterning in adolescents[30] suggest that measurement of both total and regional

fatness may be informative, especially for adolescent health screenings.

The magnitude and statistical significance of the odds for uppermost quintile membership at increasing levels of body fat, relative to the leanest group, were somewhat greater for the BP variables than for the lipoprotein variables (Figures 1 and 2). This difference in adjusted odds ratios between the BP variables and lipoprotein ratio variables may best be explained by differences in uppermost quintile membership rates within the leanest reference groups (<10% fat, males; <20% fat, females). In males, for example, for whom the difference in odds ratios between the BP variables and lipoprotein ratio variables at each level of body fat was most evident, the reference body-fat group was significantly protected against elevated BP but not against elevated lipoprotein ratios (Table 2). The trend observed in males—for the lowest risk for elevated lipoprotein ratios to occur in the 10% to 14.9% fat group—somewhat parallels the trend for higher mortality rates in the lowest-weight-for-height group than in the middle-weight-for-height group observed in adult participants of the Framingham Study.[31] However, none of the middle percent-fat groups was significantly more protected against elevated total cholesterol or lipoprotein ratios than was the leanest group (Figure 1).

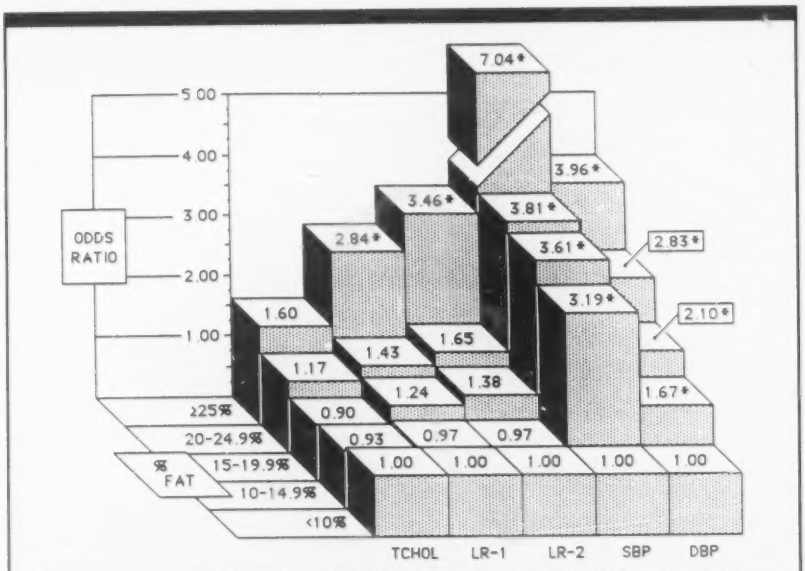


FIGURE 1—Adjusted estimates of relative odds for elevated total cholesterol, serum lipoprotein ratios, and blood pressure in males (*95% CI does not include 1.00). Abbreviations are the same as in Table 2.

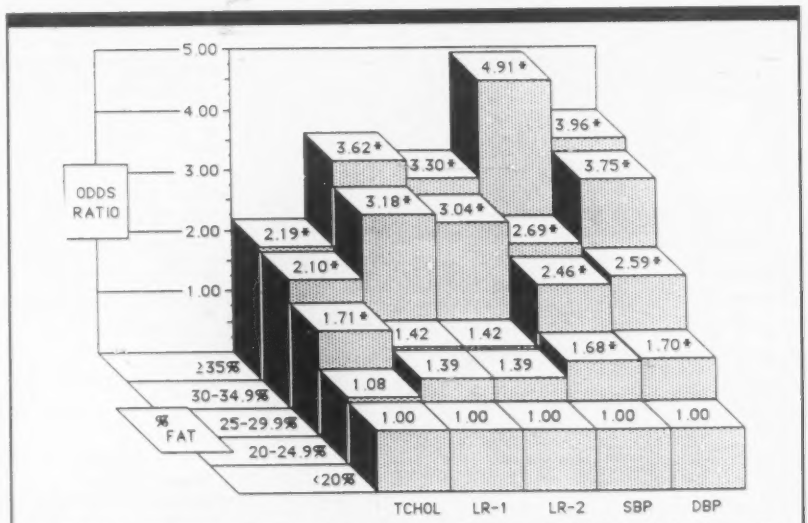


FIGURE 2—Adjusted estimates of relative odds for elevated total cholesterol, serum lipoprotein ratios, and blood pressure in females (*95% CI does not include 1.00). Abbreviations are the same as in Table 2.

In an analysis of national survey data, 6- to 11-year-old children with triceps skinfold thicknesses at or above the 85th percentile were found to be 2.6 and 1.6 times as likely as their leaner peers to have elevated systolic BP and diastolic BP, respectively.[1] Similarly, in 12- to 17-year-old youth, the risks for elevated systolic BP and diastolic BP were 4.3 and 6.1 times higher, respectively in youth having skinfold thicknesses equal to or exceeding the 85th percentile. Although the odds ratio estimates from the present study compare reasonably well with those reported previously for adolescents, they are somewhat higher than those reported for children.[1]

Several differences between the present study and the national survey analysis may account for the discrepancies in the reported relationships between body fatness and elevated BP. The 85th percentile for triceps skinfold thickness was used as an index of body fat in the previous study,[1] whereas in the present study, percent body fat was estimated from the STSS. Also, the previous study[1] referenced youth of both sexes with skinfold thicknesses equal to or exceeding the 85th percentile against those with skinfold thicknesses below the 85th percentile; however, in the present study, we contrasted four separate gender-specific percent fat groups at 5% intervals against the leanest males (<10% fat) and females (<20% fat), as appropriate. In addition, elevated BP was defined with a more stringent cutoff point (≥95th percentile) in the previous study[1] than was used in this study (≥80th percentile). Finally, in the present study, potential confounding by fat patterning was statistically controlled by using a contrast for the quintile level of $\log_n$ (subscapular/triceps); seasonality, but not fat pattern, was controlled in the previous investigation.[1]

Because a relative index of body fatness corresponding to the 85th percentile for triceps skinfold thickness was used in the earlier study,[1] it is not surprising that substantially lower estimates of risk were observed in the younger children than in the adolescents. In the present study, for example, the STSS corresponding to the

85th percentile results in percent-fat estimates of 14.8% in a 6-year-old Black male and 34.3% in a 12-year-old Black male. This age-related change in percent-fat estimates is due to the greater STSS at the 85th percentile at age 12 than at age 6 (54.8 vs 17.3 mm), and not to inaccuracies introduced by the conversion of the STSS to percent fat in children of different ages. The age-adjusted, gender- and race-specific equations used to estimate density from the STSS in the present study correct for the age-related change in the skinfold–to–density relationship in youth.[22] Moreover, the age-adjusted gender-specific equations used to convert the predicted densities to percent fat were derived from a multicomponent body composition model that accounts for age-related differences in the water and bone mineral content of the fat-free body that are associated with growth and development.[21,22]

Individual skinfold thicknesses may be inaccurate indicators of body fatness, inasmuch as a constant skin-fold thickness corresponds to a different body density and percent body fat at different ages[22] and in different races.[23] In the present study, with percent body fat calculated with an age- and race-specific approach, logistic regression coefficients were all nonsignificant for race ($P > .05$), and age was significant only for BP in males. With respect to age, statistical control for a constant relative rank of trunkal fat patterning may have exerted a stronger effect in the older adolescents than in the younger children owing to the pubescence-related increase in the absolute amount of centrally located body fat.[24,30] Thus, the body fat standards recommended in this study (≥25% for males, ≥30% for females) appear to be generally applicable throughout the 5- to 18-year-old biracial sample reported herein. However, possible biases due to subject selection, nonresponse, and geographic location may limit the generalizability of these standards, and point to the need for further validation of these standards in national probability-based samples.

On the basis of the present findings, we conclude that fatness levels at or above 25% in males and 30% in females are indicative of increased risk for elevated BP and lipoprotein ratios in White and Black children and adolescents. These new body fatness standards are applicable not only to epidemiologic studies of the prevalence and incidence of obesity in children and adolescents, but also to pediatric health screenings and health-related tests of youth fitness.[12,13] Previous studies have defined childhood obesity with age- and sex-specific values for triceps skinfold thickness[1] or weight for height[32] that correspond to critical positions in a reference population. The use of norm-referenced standards to distinguish between obese and nonobese persons is limited by the fixation of obesity prevalence to a certain percentage of a population[33] at a particular point in time and by the considerable variation in body fatness at a constant percentile of skinfolds or weight for height.[7,8] Our findings support the concept of body fatness standards as significant predictors of CVD risk factors in children and youth. Furthermore, this study offers an empirically determined alternative to current normative-based definitions of childhood obesity. Future studies in longitudinal cohorts that make direct comparisons between critical percent fat and skinfold thickness percentile cutoff points are needed to assess the specificity and sensitivity of the two screening tools. □

Acknowledgments

This research was supported by USPHS grant HL 38844 from the National Heart, Lung, and Blood Institute.

We are sincerely grateful to the children and parents of Bogalusa, La, whose participation and cooperation made this study possible. We also wish to thank Michael J. Hewitt, PhD, for his work on the illustrations.

This paper was presented in part at a symposium on youth fitness during a meeting of the Southwest Chapter of the American College of Sports Medicine in San Diego, Calif, December 1990.

References

1. Gortmaker SL, Dietz WH, Sobol AM, Wehler CA. Increasing pediatric obesity in the United States. *Am J Dis Child.* 1987; 141:535–540.
2. Aristimuño GG, Foster TA, Voors AW, Srinivasan SR, Berenson GS. Influence of persistent obesity in children on cardiovascular risk factors: the Bogalusa Heart Study. *Circulation.* 1984;69:895–904.
3. Berenson GS, Webber LS, Srinivasan SR, Voors AW, Harsha DW, Dalferes ER. Biochemical and anthropometric determinants of serum β- and pre-β-lipoproteins in children: the Bogalusa Heart Study. *Arteriosclerosis.* 1982;2:325–334.
4. Voors AW, Harsha DW, Webber LS, Radhakrishnamurthy B, Srinivasan SR, Berenson GS. Clustering of anthropometric parameters, glucose tolerance, and serum lipids in children with high and low β- and pre-β-lipoproteins in children: the Bogalusa Heart Study. *Arteriosclerosis.* 1982; 2:346–355.
5. Smoak CG, Burke GL, Freedman DS, Webber LS, Berenson GS. Relation of obesity to clustering of cardiovascular disease risk factors in children and young adults: the Bogalusa Heart Study. *Am J Epidemiol.* 1987;125:364–372.
6. Gutin B, Basch C, Shea S, et al. Blood pressure, fitness and fatness in 5- and 6-year-old children. *JAMA.* 1990;264: 1123–1127.
7. Lohman TG, Going SB, Slaughter MH, Boileau RA. Concept of chemical immaturity in body composition estimates: implications for estimating the prevalence of obesity in childhood and youth. *Am J Hum Biol.* 1989;1:201–204.
8. Going S, Williams D. Understanding fitness standards. *J Phys Educ Recreation and Dance.* 1989;60:34–38.
9. Shear CL, Freedman DS, Burke GL, Harsha DW, Berenson GS. Body fat patterning and blood pressure in children and young adults: the Bogalusa Heart Study. *Hypertension.* 1987;9:236–244.
10. Freedman DS, Srinivasan SR, Harsha DW, Webber LS, Berenson GS. Relation of body fat patterning to lipid and lipoprotein concentrations in children and adolescents: the Bogalusa Heart Study. *Am J Clin Nutr.* 1989;50:930–939.
11. Gillum RF. The association of the ratio of waist to hip girth with blood pressure, serum cholesterol and serum uric acid in children and youths aged 6–17 years. *J Chron Dis.* 1987;40:413–420.
12. *Fitnessgram user's manual.* Dallas, Tex: Institute for Aerobics Research; 1987.
13. McSwegin P, Pemberton C, Petray C, Going S. *Physical best: The AAHPERD guide to physical fitness education and assessment.* Reston, Va: American Alliance for Health, Physical Education, Recreation and Dance; 1989.
14. Berenson GS, McMahan CA, Voors AW, et al. *Cardiovascular Risk Factors in Children: The Early Natural History of Atherosclerosis and Essential Hypertension.* New York, NY: Oxford University Press; 1980.
15. Foster TA, Voors AW, Webber LS, Frerichs RR, Berenson GS. Anthropometric and maturation measurements of children ages 5–14 years, in a biracial community—the Bogalusa Heart Study. *Am J Clin Nutr.* 1977;30:582–591.
16. Foster TA, Berenson GS: Measurement error and reliability in four pediatric cross-sectional surveys of cardiovascular disease risk factor variables—the Bogalusa Heart Study. *J Chron Dis.* 1987;40:13–21.
17. Harsha DW, Frerichs RR, Berenson GS. Densitometry and anthropometry of black and white children. *Hum Biol.* 1978;50: 261–280.
18. Lussier L, Buskirk ER. Effects of an endurance training regimen on the assessment of work capacity in prepubertal children. In: Milvy P, ed. *Annals of the New York Academy of Science.* New York, NY: 1977:734–747.
19. Mukherjee D, Roche AF. The estimation of percent body fat, body density and total body fat by maximum R^2 regression equations. *Hum Biol.* 1984;56:79–109.
20. Slaughter MH, Lohman TG, Boileau RA, Van Loan M, Horswill CA, Wilmore JH. Influence of maturation on relationship of skinfolds to body density: a cross-sectional study. *Hum Biol.* 1984;56:681–689.

21. Lohman TG. Assessment of body composition in children. *Pediatr Exec Sci.* 1989;1: 19–30.
22. Lohman TG. Applicability of body composition techniques and constants for children and youth. In: Pandolf KB, ed. *Exercise and Sport Sciences Reviews.* Vol. 14. New York, NY: Macmillan Publishing Co Inc; 1986:325–357.
23. Slaughter MH, Lohman TG, Boileau RA, et al. Skinfold equations for estimation of body fatness in children and youth. *Hum Biol.* 1988;60:709–723.
24. Baumgartner RN, Roche AF. Tracking of fat pattern indices in childhood: the Melbourne Growth Study. *Hum Biol.* 1988; 60:549–567.
25. *Manual of Laboratory Operation, Lipid Research Clinics Program.* Vol. 1. Bethesda, Md: National Institutes of Health; 1974. US Dept of Health, Education, and Welfare publication NIH 75-628.
26. Srinivasan SR, Berenson GS. Serum lipoproteins in children and methods for study. In: Lewis LA, ed. *CRC Handbook of Electrophoresis. Vol. 3. Lipoprotein Methodology and Human Studies.* Boca Raton, Fla: CRC Press; 1983:185–204.
27. Shear CL, Burke GL, Freedman DS, Webber LS, Berenson GS. Designation of children with high blood pressure—considerations on percentile cut points and subsequent high blood pressure: the Bogalusa Heart Study. *Am J Epidemiol.* 1987; 125:73–84.
28. Schlesselman JJ. *Case-control Studies: Design, Conduct, Analysis.* New York, NY: Oxford University Press; 1982.
29. Freedman DS, Srinivasan SR, Burke GL, et al. Relation of body fat distribution to hyperinsulinemia in children and adolescents: the Bogalusa Heart Study. *Am J Clin Nutr.* 1987;46:403–410.
30. Williams DP, Going SB, Lohman TG, et al. Trunkal fat patterning and its relation to cardiovascular disease (CVD) risk factors in children and adolescents. *Med Sci Sports Exerc.* 1991;23:S51.
31. Sorlie P, Gordon T, Kannel WB. Body build and mortality. *JAMA.* 1980;243: 1828–1831.
32. Newens EM, Goldstein H. Height, weight and the assessment of obesity in children. *Br J Prev Soc Med.* 1971;26:33–39.
33. Weil WB. Obesity in childhood. In: Filer LJ, Lauer RM, eds. Children's blood pressure report of the eighty-eighth Ross conference on pediatric research. Columbus, Ohio: Ross Laboratories; 1985:103–108.

Deadline for Student Papers on Injury Is March 31, 1992

The Southern California Injury Prevention Research Center is pleased to announce a competition for the best student papers in 1991 on injury. Papers must focus on public health aspects of intentional injury (homicide, suicide, sexual assault, battering, child abuse, etc.) or unintentional injury (motor vehicle crashes, falls, drowning, burns, etc.). Priority will be given to original scientific contributions that focus on injuries in ethnic or racial minorities, the poor, and other traditionally underserved populations.

Three awards are anticipated in the amounts of $500, $250, $150 for manuscripts that demonstrate original methods in injury research, analyze new data, or emphasize findings with countermeasure potential. The award will be made directly to the student in recognition of his or her contribution to injury research.

Although papers published in 1991 will be considered, the submission of unpublished manuscripts is strongly encouraged. Papers should be submitted by the student or by an academic adviser on behalf of the student and with his or her permission. If unpublished, the manuscript (including references and tables) should not exceed 30 double-spaced pages. In addition to the manuscript, a letter from the student's adviser should be submitted which verifies that the work and writing were conducted by the student who was enrolled in graduate training during 1991. Full acknowledgment of the student's primary role in the work should be reflected through his or her first authorship of the paper. All submissions must include student's name, address, and daytime telephone number. Manuscripts should be prepared according to the *American Journal of Public Health* style guidelines and submitted in triplicate to Jess F. Kraus, PhD, Southern California Injury Prevention Research Center, UCLA School of Public Health, 10833 Le Conte Avenue, Los Angeles, CA 90024-1772. The deadline for submissions, which was omitted from the previous announcement in the January issue of the *American Journal of Public Health,* is March 31, 1992.

Inquiries may be directed to Renee Goetz, MSW, Administrative Coordinator, at the same UCLA address (telephone 213/206-4115).

Childhood Chronic Illness: Prevalence, Severity, and Impact

ABSTRACT

Background. Using data from the 1988 National Health Interview Survey, this article presents national estimates of the prevalence and impact of childhood chronic conditions.

Methods. Proxy responses to a checklist of child health conditions administered for 17 110 children under 18 years of age were used. Conditions were classified as chronic if they were first noticed more than 3 months prior to the interview or if they were the type that would ordinarily be of extended duration, such as arthritis.

Results. An estimated 31% of children were affected by chronic conditions. Among these children, highly prevalent conditions included respiratory allergies 9.7 per 100, repeated ear infections 8.3 per 100 and asthma 4.3 per 100. These children can be divided into three groups: 66% with mild conditions that result in little or no bother or activity limitation; 29% with conditions of moderate severity that result in some bother or limitation of activity, but not both; and 5% with severe conditions that cause frequent bother and limitation of activity. The 5% with severe conditions accounted for 19% of physician contacts and 33% of hospital days related to chronic illness.

Conclusions. Childhood chronic conditions have highly variable impacts on children's activities and use of health care. (*Am J Public Health.* 1992;82:364–371)

Paul W. Newacheck, DrPH, and William R. Taylor, MD

Introduction

Childhood illness has changed dramatically in the United States during this century. Infectious diseases that claimed the lives of thousands of children during the early 20th century have been largely eliminated through improvements in infectious disease control, sanitation, housing, and medical care. Mortality rates have fallen from 870 per 100 000 children aged 1 to 14 years in 1900 to 33 per 100 000 children in 1987.[1] But although this trend demonstrates progressive improvements along one critical dimension of child health, reductions in the prevalence of childhood chronic illness have not been nearly as dramatic.[2]

True, the prevalence of some childhood chronic conditions, such as polio and rheumatic fever, has declined markedly in the second half of this century. However, dramatic medical advances in the last few decades have resulted in many chronically ill children, who previously would have died much earlier from their illnesses, now surviving into adulthood. For example, there is evidence of a sevenfold increase in survival to age 21 among children with cystic fibrosis and of twofold or greater increases in survival for children with spina bifida, leukemia, and congenital heart disease.[3,4] Moreover, evidence from other quarters, albeit more controversial, suggests that the prevalence of non-life-threatening chronic conditions may have increased in recent years. According to national household survey data, the prevalence of asthma severe enough to limit a child's activities increased by 65% during the 1970s. Similar increases were reported for hearing impairments and a variety of mental and nervous system disorders.[5]

Despite concern over these trends, knowledge of the prevalence and distribution of childhood chronic illness is limited.[6] Indeed, there are not even accepted criteria for defining childhood chronic illness. Partly because such common definitional criteria are lacking, the research community has generated widely differing estimates of the prevalence of childhood chronic conditions. Additional variation results from the different methods used in collecting prevalence data. Some studies are population based; others are clinic based. Some rely on parental reports; others use medical records. Some include only chronic physical conditions; others include physical and mental conditions. As a result of these different methods, published estimates of the proportion of children with one or more chronic illnesses range from less than 5% to more than 30%.[7–13]

The present study provides estimates of the prevalence of childhood chronic conditions from a large population-based sample of children included in the National Health Interview Survey (NHIS) on Child Health.[14] Using this survey to derive prevalence estimates offers a number of advantages. First, a population-based sample provides for more representative results than could be obtained from a sample of patients appearing at clinics, hospitals, or other treatment centers. Second, a nationally representative sample eliminates geographic biases that might result

Paul Newacheck is with the Institute for Health Policy Studies, University of California, San Francisco. William Taylor is with the Office of Program Planning and Evaluation, Centers for Disease Control, Atlanta.

Requests for reprints should be sent to Paul W. Newacheck, DrPH, Institute for Health Policy Studies, 1388 Sutter Street, 11th floor, San Francisco, CA 94109.

This paper was submitted to the Journal February 21, 1991, and accepted with revisions August 23, 1991.

from localized small-area samples. Third, the large sample size (n = 17 110) permits more precise prevalence estimates than could be obtained from smaller samples.

Data from the survey are used in this paper to examine the national prevalence of childhood chronic illness. We present prevalence estimates for 19 different categories of chronic conditions and estimates of the number of children with multiple chronic conditions. We describe how prevalence varies according to age, gender, and race. We examine the impact of chronic conditions on children's activity levels and show how childhood chronic conditions influence children's use of health services. Finally, we show how prevalence estimates change when different severity thresholds are applied.

Methods

The NHIS is a continuing nationwide survey conducted by the Bureau of the Census for the National Center for Health Statistics (NCHS). Its purpose is to collect information on the health status and use of health care by the US population. The sample is designed to be representative of the US civilian noninstitutional population.[14]

In 1988, a special supplemental questionnaire on child health was included in the survey. This questionnaire, similar to one used in 1981, covered a range of topics relevant to child health, including child care, pregnancy and birth, behavioral problems, developmental problems, other child health problems, and use of health services. One child in each household was selected at random to be the subject of the supplemental interview. An adult member of the household knowledgeable about the child's health served as the respondent. In about 90% of the cases, the respondent was a parent of the sample child, usually the mother. Interviews were completed for 91% of all eligible children, resulting in a sample of 17 110 children under 18 years old.[15]

Of particular importance was the inclusion in the supplement of a detailed checklist of childhood health conditions. The checklist was oriented toward recurrent or chronic conditions. Respondents were read this extensive list and asked whether the sample child had any of the included conditions during the previous 12 months. From these results, annual prevalence estimates can be derived for a variety of childhood health problems as diverse as repeated ear infections, cerebral palsy, and juvenile arthritis—but excluding minor acute illnesses, such as colds and the flu. In part because of past problems of underreporting, cancers and mental health problems without physical manifestations were excluded from the checklist.

Data on the impact of checklist conditions were also collected in the supplement. Each time a condition was reported as present during the previous year, a set of follow-up questions was asked about the frequency and amount of bother caused by the condition; the days spent ill in bed; the days missed from school; and the resultant use of medications, physician services, and hospital care. Together, these data permitted analysis of the impact of chronic conditions on children's activity levels and health care use patterns. The questions and probes used to identify the impact of chronic conditions are reproduced in Appendix A.

Data on school absences and days in bed were not collected for children experiencing hearing, vision, or speech problems or eczema and skin allergies on the assumption that these conditions result in few, if any, bed days or school absences. A follow-up question on limitations in play and sports activities was included in the supplement but was asked only for children with certain chronic conditions. However, the core questionnaire contained questions that were asked for all children regarding limitation of activity due to chronic conditions, and those results are used here. With the results from those questions, children were classified as limited in their activities if they were reported as being limited in or unable to conduct age-appropriate activities, such as school or play.

We asked a panel of eight pediatricians to review the 76 checklist conditions to determine whether any should be excluded from our analysis of chronic conditions. When a majority considered a condition not to be chronic, that condition was eliminated. A total of eight conditions were eliminated in this fashion. We then classified the remaining conditions according to chronicity using conventions similar to those developed by the NCHS's Division of Health Interview Statistics. Specifically, a condition was considered chronic if (1) the respondent indicated it was first noticed more than 3 months before the interview date, or (2) it was a type of condition that ordinarily has a duration of more than 3 months.[16] Examples of conditions that are considered chronic regardless of their time of onset include diabetes, heart conditions, and arthritis.

Because no severity criteria are used, this approach to defining chronic illness leads to the inclusion of a large number of conditions and thus provides an upper-bound estimate of the size of the chronically ill child population. We show how prevalence estimates diminish when we use more restrictive definitions that take into account the impact or severity of reported conditions.

To enable a manageable analysis and presentation, we grouped conditions into 19 broader categories, including 5 impairment groupings and 14 disease groupings. This categorization scheme is based on grouping conditions that are clinically related while taking into account constraints due to sample size limitations. The disease categorization includes conditions with active pathology whereas the impairment categorization includes conditions with stable pathology. The overall categorization scheme and condition groupings are illustrated in Appendix B.

The actual number of sample cases is presented in the first column of Table 1 and Table 2. All other results are weighted to reflect national population estimates. The weights were provided by NCHS and are equal to the inverse of the sampling probability for each sample child. Standard errors were calculated using software designed to take into account the complex sample design of the NHIS.[17,18]

Results

Prevalence of Childhood Chronic Conditions

As shown in Table 1, an estimated 31% of children under 18 years of age, or almost 20 million children nationwide, were reported to have one or more chronic conditions in 1988. Although most chronically ill children were affected by a single condition, a significant minority were afflicted by multiple chronic conditions. Among children with chronic conditions, 70% were reported to have one checklist condition, 21% had two, and 9% had three or more. It is likely that these estimates understate the true size of the child population with single and multiple conditions, given that not all childhood chronic conditions—particularly those related to mental health—were included in the checklist.

As shown in Table 2, the most commonly reported childhood chronic conditions included respiratory allergies (including hay fever) and frequent or repeated ear infections. Respiratory aller-

TABLE 1—Prevalence of Chronic Conditions among Children under 18 Years of Age, 1988.

Number of Conditions	N	Cases per 1000	Std Error	Cases in 1000s
Children with one or more chronic conditions	5332	307.6	6.8	19 556
Children with one condition	3707	214.3	5.0	13 261
Children with two conditions	1145	64.8	2.5	4121
Children with three or more conditions	480	28.5	1.9	1813

Source—Original tabulations of the 1988 National Health Interview Survey.

gies affected 97 of every thousand children and 83 of every thousand children experienced repeated ear infections. Other relatively common conditions included asthma, eczema and skin allergies, frequent or severe headaches, and speech defects. Each of these had a prevalence of 25 per 1000 or more. Conditions of low prevalence included diabetes, sickle cell disease, and cerebral palsy, each of which affected fewer than two of every thousand children in the United States.

Prevalence of childhood chronic conditions is shown in Table 3 to vary by demographic status. Although chronic conditions appear slightly more prevalent among older children (315 per 1000) than among younger children (302 per 1000), the difference is not statistically significant. However, chronic conditions were more prevalent among boys (326 per 1000) than among girls (288 per 1000), and among White children (324 per 1000) than among Black children (247 per 1000).

Impact of Childhood Chronic Conditions on Children's Health

The questionnaire included several measures of the impact of chronic conditions on children's health. Included among these was both the frequency and the amount of bother caused by the condition. On average, childhood chronic conditions were reported to never bother children 25% of the time, bother them once in a while 53% of the time, and bother them often or all the time 22% of the time (Table 4). The amount of bother caused by childhood chronic conditions ranged, on average, from a great deal of bother (25%) to very little or some bother (50%) to no bother at all (25%).

The survey also collected information on limitation in usual activities and on average annual bed days and school absences (Table 5). Overall, 13% of children with one or more chronic conditions were classified as limited in their usual activities. Children with one or more chronic conditions experienced an average of slightly more than 2 bed days and 3 lost school days in the previous year. These bed days and school absences are in addition to those resulting from acute conditions such as colds, the flu, and minor injuries. Hence, they represent the added burden of chronic illness for these children.

The average numbers of bed days and school absences presented in Table 5 mask large differences among individual children even with the same condition. In fact, bed days and lost school days were quite unevenly distributed among children with chronic conditions. Most children experienced no bed days or school absences in 1988 that could be attributed to chronic conditions; it was just a small segment of the chronically ill child population that accounted for the majority of bed days and lost school days. Specifically, 72% of children with chronic conditions spent no days confined to bed in the previous year, and 59% had no school absences attributable to chronic illness. At the other end of the spectrum, 8% of children with chronic conditions spent 7 or more days confined to bed in the previous year and accounted for 74% of all bed days caused by chronic illness. Similarly, 12% of the population with chronic conditions had 7 or more school absence days and accounted for 73% of all school absences related to chronic health problems. Hence, a small segment of the chronically ill population disproportionately shoulders the burden of illness when that illness is measured in terms of bed days and school absences.

Impact of Childhood Chronic Conditions on Children's Use of Services

Use of medical services for children with chronic conditions is shown in Table 6. In 1988, 63% of all children with chronic conditions were reported by their parents to have used medications (excluding vitamins) prescribed or recommended by a doctor. In addition, 4% of children were hospitalized to treat their chronic health problems sometime during the year prior to the interview. Nationally, an estimated 690 000 children were hospitalized for a total of nearly 7.2 million days for treatment of chronic conditions.

Children averaged slightly under five physician contacts annually for the treatment of chronic conditions (excluding those occurring during inpatient hospital episodes). Nationally, chronically ill children were estimated to have made 92 million ambulatory physician contacts in 1988, excluding visits for routine preventive care and services obtained in treating acute conditions. Hence, these results illustrate the added health care burden of childhood chronic illness.

Like bed days and school absences, physician services were used at markedly varying rates by individual chronically ill children. Some children made very high use of those services whereas many had no physician contacts during the previous year. Specifically, 30% of children with chronic conditions had no physician services related to the care of chronic illnesses in 1988. In contrast, 10% had 10 or more contacts and accounted for 63% of all ambulatory physician services for chronic illness. Hence, a small segment of the chronically ill child population appears to account for the majority of ambulatory care services.

The Uneven Burden of Childhood Chronic Illness

The data presented above suggest that, although chronic conditions are quite common among children, a comparatively small number of children are severely affected by them. For example, only one in eight children with chronic conditions had any limitation on his or her usual activities. Similarly, most children experienced only occasional bother or no bother at all from their chronic conditions. Further analysis indicates that 66% of children with chronic conditions were never or only occasionally bothered by their conditions and had no limitations on their usual activities. These data suggest that two thirds of all children with chronic conditions are mildly affected by their conditions and experience little or no adverse effects on their activities.

At the opposite end of the severity spectrum are children who were bothered by their conditions often or all of the time *and* limited in their usual activities

TABLE 2—Prevalence of Specified Chronic Conditions among Children under 18 Years of Age, 1988

Condition	N	Cases per 1000	Std Error	Cases in 1000s
All children with chronic conditions	5332	307.6	6.8	19 556
Impairments				
Musculoskeletal impairments	232	15.2	1.5	967
Deafness and hearing loss	241	15.3	1.1	975
Blindness and vision impairment	220	12.7	1.2	810
Speech defects	388	26.2	1.9	1666
Cerebral palsy	34	1.8	0.4	112
Diseases				
Diabetes	25	1.0	0.2	64
Sickle cell disease	26	1.2	0.3	74
Anemia	139	8.8	1.0	557
Asthma	747	42.5	1.8	2700
Respiratory allergies	1746	96.8	3.7	6155
Eczema and skin allergies	557	32.9	1.9	2088
Epilepsy and seizures	34	2.4	0.5	151
Arthritis	79	4.6	0.6	290
Heart disease	272	15.2	1.2	965
Frequent or repeated ear infections	1480	83.4	2.7	5304
Frequent diarrhea/bowel trouble	308	17.1	1.4	1085
Digestive allergies	398	22.3	1.4	1419
Frequent or severe headaches	419	25.3	1.5	1606
Other	344	19.8	1.2	1256

Source—Original tabulations of the 1988 National Health Interview Survey.

TABLE 3—Prevalence of Childhood Chronic Conditions in Cases per 1000 by Age, Gender, and Race, 1988

Condition	Age		Gender		Race	
	Under 10 Years	10 to 17 Years	Boys	Girls	White	Black
All children with chronic conditions	302.2	315.0	326.2	288.2	324.4	246.5
Impairments						
Musculoskeletal impairments	10.9	20.9	16.7	13.6	15.9	10.5
Deafness and hearing loss	14.1	17.0	18.3	12.3	17.7	6.0
Blindness and vision impairment	10.3	16.0	11.4	14.2	13.9	8.4
Speech defects	31.6	18.9	35.3	16.7	25.9	33.5
Cerebral palsy	2.2	1.2[a]	2.0	1.5[a]	1.9	0.5[a]
Diseases						
Diabetes	0.6[a]	1.5[a]	1.5	0.5[a]	1.2	0.1
Sickle cell disease	1.3[a]	0.9[a]	0.9[a]	1.4	0.1[a]	6.9
Anemia	11.0	5.8	8.4	9.1	8.6	10.0
Asthma	39.3	46.8	50.7	33.9	41.0	51.4
Respiratory allergies	71.8	130.3	106.5	86.7	106.1	53.2
Eczema and skin allergies	31.1	35.2	30.1	35.8	34.3	22.8
Epilepsy and seizures	1.7[a]	3.3	1.7[a]	3.1	2.5	2.3[a]
Arthritis	1.5[a]	8.7	4.2	4.9	4.5	5.3[a]
Heart disease	13.6	17.4	16.4	13.9	17.3	7.5
Frequent or repeated ear infections	120.6	33.6	88.5	78.1	91.6	53.2
Frequent diarrhea/bowel trouble	22.6	9.6	18.1	15.9	17.8	13.6
Digestive allergies	23.2	21.1	25.6	18.9	24.5	9.7
Frequent or severe headaches	9.9	45.8	22.8	27.9	26.5	21.9
Other	12.1	30.0	19.3	20.3	21.9	9.9

Source—Original tabulations of the 1988 National Health Interview Survey.
[a]standard error exceeds 30% of estimate value.

by chronic illness. This severe group is much smaller, including only 5% of children with chronic conditions. In between these extremes are the remaining 29% of children with chronic conditions. This group includes children who experienced more than occasional bother or limitation of activity, but not both (see Figure 1). Expressed another way, 20% of U.S. children experienced mild chronic conditions, 9% experienced chronic conditions of moderate severity, and only 2% of children experienced severe chronic conditions in 1988.

Children in the severe group spent an average of 10 days in bed and missed 11 days from school annually. Despite their small numbers, they account for 27% of all bed days and 24% of all school absences related to chronic conditions. Their use of medical services was also extraordinary.

TABLE 4—Percent Distribution of Childhood Chronic Conditions by Frequency and Amount of Bother, 1988

Condition	Frequency of Bother			Amount of Bother		
	Never	Once in Awhile	Often or All the Time	Never	Very Little or Some	Great Deal
All conditions	24.7	53.0	22.3	24.7	50.1	25.2
Impairments						
Musculoskeletal impairments	48.6	33.2	18.2	48.6	36.4	15.0
Deafness and hearing loss	57.0	23.8	19.2	57.2	25.5	17.3
Blindness and vision impairment	66.9	15.7	17.4	67.3	23.1	9.6
Speech defects	72.8	16.1	11.1	72.8	20.0	7.2
Cerebral palsy	45.4	30.2	24.4[a]	45.4	33.7	20.8[a]
Diseases						
Diabetes	46.7	25.8[a]	27.5[a]	46.7	37.9[a]	15.4[a]
Sickle cell disease	79.2	12.2[a]	8.6[a]	79.2	17.9[a]	2.9[a]
Anemia	71.9	21.7	6.4[a]	71.9	24.6	3.5[a]
Asthma	10.3	67.6	22.1	10.3	55.3	34.4
Respiratory allergies	9.2	70.7	20.1	9.3	71.2	19.5
Eczema and skin allergies	29.0	51.1	19.9	29.1	56.3	14.6
Epilepsy and seizures	44.1	40.8	15.2[a]	44.1	36.9	19.1[a]
Arthritis	18.5[a]	47.0	34.5	18.5[a]	40.4	41.1
Heart disease	74.5	15.7	9.8	74.5	15.5	10.0
Frequent or repeated ear infections	5.3	63.4	31.3	5.3	55.9	38.7
Frequent diarrhea/bowel trouble	16.1	56.1	27.8	16.1	60.6	23.3
Digestive allergies	30.9	54.0	15.1	30.9	49.7	19.4
Frequent or severe headaches	0.8[a]	63.4	35.8	0.8[a]	41.9	57.3
Other	36.1	39.5	24.3	36.1	36.1	27.8

Source—Original tabulations of the 1988 National Health Interview Survey.
[a]standard error exceeds 30% of estimate value.

TABLE 5—Impact of Childhood Chronic Conditions on Children's Activity Levels, 1988

Condition	Percent Limited in Usual Activities Due to Chronic Conditions	Average Annual Bed Days Due to Chronic Conditions	Average Annual School Absence Days Due to Chronic Conditions
All children with chronic conditions	13.3	2.2	3.1
Impairments			
Musculoskeletal impairments	39.6	1.4[a]	2.3[a]
Deafness and hearing loss	35.5	—[b]	—[b]
Blindness and vision impairment	19.4	—[b]	—[b]
Speech defects	33.1	—[b]	—[b]
Cerebral palsy	89.2	1.7[a]	2.3[a]
Diseases			
Diabetes	30.0[a]	3.6[a]	3.1[a]
Sickle cell disease	15.2[a]	0.2[a]	0.2[a]
Anemia	14.7	0.8[a]	0.8[a]
Asthma	28.6	2.8	4.6
Respiratory allergies	12.3	0.8	1.4
Eczema and skin allergies	10.5	—[b]	—[b]
Epilepsy and seizures	58.0	3.0[a]	3.4[a]
Arthritis	24.6	3.3[a]	3.2[a]
Heart disease	21.6	3.2[a]	2.2[a]
Frequent or repeated ear infections	9.1	1.7	2.8
Frequent diarrhea/bowel trouble	16.7	0.9[a]	1.4
Digestive allergies	15.7	0.4	1.1
Frequent or severe headaches	16.9	2.9	3.3
Other	24.7	2.8	3.4

Source—Original tabulations of the 1988 National Health Interview Survey.
[a]standard error exceeds 30% of estimate value.
[b]information on bed days and school absences was not obtained for these conditions.

Children in this group had an average of 16 physician contacts annually, and 16% were hospitalized during the previous year. Together they account for 19% of all physician contacts and 33% of all hospital days resulting from chronic conditions. In addition, 85% of children in this group received physician-recommended medications.

As one might expect, the diagnostic case mix differs substantially at each extreme of the severity spectrum. Children with mild chronic conditions were much more likely than children with severe chronic conditions to have respiratory allergies and repeated ear infections. In contrast, children in the severe group were more likely than children in the mild group to have musculoskeletal impairments, hearing and speech impairments, cerebral palsy, diabetes, asthma, epilepsy, and arthritis. In addition, children in the severe group had an average of 2.6 conditions, compared with an average of only 1.3 conditions per child in the mild group.

Discussion

In reviewing the NHIS results presented here, the reader should keep in mind that the results are based on household interviews, not on diagnostic tests or medical examinations. These results represent the mother's or other adult's *perception* of the child's health. Respondents are more likely to report conditions that have some impact on a child's activities.

Similarly, conditions associated with recent medical encounters are more likely to be accurately reported in household interviews. In contrast, conditions with little functional impact are often underreported, as are conditions that are associated with embarrassment or stigma.[19,20]

Although NCHS has undertaken validation studies, none of these studies has focused on children. Comparison of NHIS interview data from the early 1960s with information derived from medical records for *adults* suggests that chronic conditions, particularly those not causing significant disability, are generally underreported in household interviews. The most recent study of NHIS interviewing methodology found that between 35% and 45% of conditions found in medical records were unreported in household interviews whereas between 8% and 12% of conditions reported in interviews were not found on medical records. Conditions most likely to be underreported in household interviews include those related to the urinary and reproductive systems, mental illness, and cancer.[19,20] The results presented here for children contain at least one instance to suggest that misreporting may occur among respondents. Specifically, 18% of child diabetics were reported as unmedicated in the last year, indicating that either parents overreport diabetes or underreport use of medications.

With these limitations in mind, results from this investigation suggest that childhood chronic conditions are quite common in the United States. Using a broad definition of chronic illness, approximately 31% of all children under 18 years of age were classified as having chronic conditions in 1988. The overall prevalence of chronic conditions was similar for adolescents and younger children. Chronic conditions were reported more frequently for boys than for girls and for Whites than for Blacks. In fact, differences in prevalence were especially large for the two racial groups; however, whether these racial differences are attributable to reporting biases or underlying differences in health cannot be discerned from this data set.

Our results indicate that children with chronic conditions experience a substantial added burden of illness. In 1988, chronically ill children tallied 37 million bed days and 41 million school absence days attributable to chronic conditions. As previously noted, they also made 92 million physician contacts and accounted for over 7 million hospital days related to the treatment of chronic conditions. Using an average daily cost of $581 for hospital care[21] and $36 for intermediate-level visits for established pediatric patients (CPT 90060),[22] we estimate the cost of hospital and physician care related to childhood chronic conditions to be approximately $7.5 billion annually. This estimate may be somewhat inflated because physician contacts as defined in the questionnaire include consultations in person or by telephone. At the same time, however, this estimate does not include charges for nonphysician professional services, medications, supplies and equipment, and other specialized services used by chronically ill children.

Our results also indicate that the bur-

TABLE 6—Impact of Childhood Chronic Conditions on Use of Medical Services, 1988

Condition	Percent Using Medications in Last Year Due to Chronic Conditions	Percent Hospitalized in Last Year Due to Chronic Conditions	Average Annual Physician Contacts Due to Chronic Condition
All children with chronic conditions	63.1	3.8	4.7
Impairments			
Musculoskeletal impairments	15.4	4.9[a]	3.1
Deafness and hearing loss	27.1	1.8[a]	2.1
Blindness and vision impairment	7.1	1.5[a]	1.5
Speech defects	1.8[a]	—[b]	4.8[a]
Cerebral palsy	23.5[a]	12.1[a]	10.3[a]
Diseases			
Diabetes	81.7	35.5[a]	7.9[a]
Sickle cell disease	22.8[a]	<0.1	2.2[a]
Anemia	42.7	0.6[a]	2.3
Asthma	84.9	7.4	4.9
Respiratory allergies	63.4	0.5[a]	2.8
Eczema and skin allergies	66.8	—[b]	1.6
Epilepsy and seizures	84.3	19.1[a]	3.7
Arthritis	32.8	5.6[a]	6.3[a]
Heart disease	15.2	7.9	2.7
Frequent or repeated ear infections	93.2	2.3	4.3
Frequent diarrhea/bowel trouble	46.0	2.7[a]	1.9
Digestive allergies	36.0	1.0[a]	2.8
Frequent or severe headaches	47.5	1.3[a]	1.6
Other	38.9	8.4	4.5

Source.—Original tabulations of the 1988 National Health Interview Survey.
[a]standard error exceeds 30% of estimate value.
[b]Information on hospitalization was not obtained for these conditions.

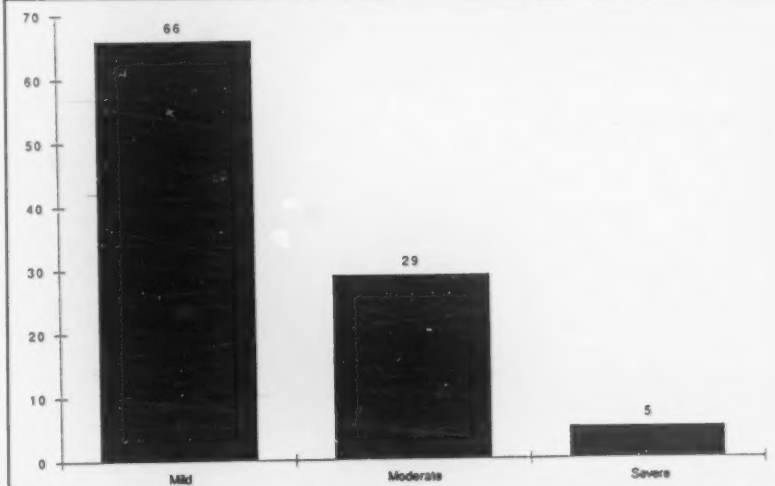


FIGURE 1—Percentage distribution of children with chronic conditions, by severity level, 1988.

den of chronic illness is not borne evenly by these children. Chronic conditions fall along a continuum that begins with those that have essentially no impact on the child and ends with those that affect the child and the family profoundly. Using this continuum approach, the 31% of US children with chronic conditions can be divided into three groups: 20% with mild conditions that result in little or no bother or activity limitations; 9% with conditions of moderate severity that result in some bother or limitation of activity, but not both; and 2% with severe conditions that cause frequent bother and limitation of activity.

In prior prevalence studies using other data sources, prevalence estimates for childhood chronic conditions have ranged from less than 5% to more than 30%. The overall prevalence estimate of 31% from the 1988 NHIS is clearly at the upper extreme of previous prevalence estimates whereas our 2% estimate for the prevalence of severe conditions is clearly at the lower extreme. This illustrates the sensitivity of prevalence estimates to the definition used.

Results from the NHIS clearly demonstrate the wide variability with which specific conditions affect the health of children. The impact of chronic conditions on a given child is influenced by their type, number, and severity. The impact is also determined by external factors, such as psychosocial support within the family, access to medical and other services, effectiveness of treatment and rehabilitative interventions, and acceptance by the community of children with chronic conditions. With a goal to diminish the impact of illness and prevent dysfunction in children, the variability in impact on health warrants further study to determine which external factors could be modified to improve health. □

Acknowledgments

This research was supported in part by Centers for Disease Control and the Federal Bureau of Maternal and Child Health.

The authors thank Deborah Peltzman for programming assistance and Helen Gonzales for help in preparing the manuscript. We also benefited from the helpful comments of Coleen Boyle, Neal Halfon, Jim Marks, Margaret McManus, and members of the Research Consortium on Chronic Illness in Childhood.

References

1. National Center for Health Statistics. *Prevention profile. Health, United States, 1989.* Hyattsville, Md: Public Health Service; 1990.
2. Hobbs N, Perrin JM, Ireys HT, eds. *Chronically Ill Children and Their Families.* San Francisco, Calif: Jossey-Bass Inc; 1985.
3. Gortmaker S, Sappenfield W. Chronic childhood disorders: prevalence and impact. *Pediatr Clin North Am.* 1984;31:3–18.
4. Gortmaker SL. Demography of chronic childhood diseases. In: Hobbs N, Perrin JM, eds. *Issues in the Care of Children with Chronic Illness.* San Francisco, Calif: Jossey-Bass Inc; 1985.
5. Newacheck PW, Budetti PP, Halfon N. Trends in activity-limiting chronic conditions among children. *Am J Public Health.* 1986;76:178–184.
6. Starfield B. The state of research on chronically ill children. In: Hobbs N, Perrin JM, eds. *Issues in the Care of Children with Chronic Illness.* San Francisco, Calif: Jossey-Bass Inc; 1985.
7. Stewart W. The unmet needs of children. *Pediatrics.* 1967;39:157–160.
8. Cadman D, Boyle M, Szatmari P, et al. Chronic illness, disability, and mental and social well-being: findings from the Ontario Child Health Study. *Pediatrics.* 1987; 79:805–813.
9. Pless I, Roghman K. Chronic illness and its consequences: observations based on three epidemiologic surveys. *J Pediatr.* 1971; 79:351–359.
10. Walker D, Gortmaker S, Weitzman M. *Chronic Illness and Psychosocial Problems among Children in Genesee County.* Cambridge, Mass: Harvard School of Public Health; August 1981.
11. Gortmaker SL, Walker DK, Weitzman M, Sobol AM. Chronic conditions, socioeconomic risks, and behavioral problems in children and adolescents. *Pediatrics.* 1990;85:267–276.
12. Newacheck PW, Halfon N, Budetti PP. Prevalence of activity-limiting chronic conditions among children based on household interviews. *J Chron Dis.* 1986;39:63–71.
13. Collins JG. Prevalence of selected chronic conditions, United States, 1983–85. *Vital and Health Statistics* (advance data). Series, No. 155. Hyattsville, Md: Public Health Service, National Center for Health Statistics; May 1988. Dept of Health and Human Services publication PHS 88–1250.
14. Adams PF, Hardy AM. Current estimates from the National Health Interview Survey: United States, 1988. *Vital and Health Statistics.* Series 10, No. 173. Hyattsville, Md: National Center for Health Statistics; October 1989.
15. Bloom B. Health insurance and medical care: health of our nation's children: United States, 1988. *Vital and Health Statistics* (advance data). Series, No. 188. Hyattsville, Md: National Center for Health Statistics; 1990.
16. US Dept of Health and Human Services. Public use data tape documentation: part III. medical coding manual and short index. *National Health Interview Survey 1987.* US Dept of Health and Human Services: Hyattsville, Md; November 1988.
17. Shah BV. *SESUDAAN: Standard Errors Program for Computing of Standardized Rates from Sample Survey Data.* Research Triangle Park, NC: Research Triangle Institute; April 1981.
18. Shah BV. *RTIFREQS: Program to Compute Weighted Frequencies, Percentages, and Their Standard Errors.* Research Triangle Park, NC: Research Triangle Institute; October 1982.
19. National Center for Health Statistics. Interview data on chronic conditions compared with information derived from medical records. Series 2, No. 23. Washington, DC: US Govt Printing Office; 1967. Public Health Service publication 1000.
20. Jabine TB. *Reporting Chronic Conditions in the National Health Interview Survey: A Review of Findings from Evaluation Studies and Methodological Tests.* Washington, DC: US Govt Printing Office; 1985. Dept of Health and Human Services P.O. 85A0046165001D.
21. American Hospital Association. *Hospital Statistics 1989–90.* Chicago, Ill; 1988.
22. Kirchner M. Where do your fees fit in? *Medical Economics.* October 1990:76–105.

APPENDIX A—Questions Used to Determine Impact of Childhood Conditions

1) During the past 12 months, did ____ (*condition/AIP*) cause ____ to miss any time from school?
2) How many days in the past 12 months did ____ miss all or part of the day?
3) During the past 12 months, did ____ (*condition/AIP*) cause ____ to stay in bed more than half of the day?
4) How many days in the past 12 months did ____ stay in bed more than half of the day?
5) During the past 12 months, about how many nights did ____ spend in the hospital because of (*condition/AIP*)?
6) During the past 12 months, about how many times did [___/anyone] see or talk to a medical doctor or assistant about this (*condition/AIP*)? (Do not count doctors seen while an overnight patient in a hospital.)
7) During the past 12 months, did this (*condition/AIP*) make it necessary for ____ to use any medicine, other than vitamins, that a doctor prescribed OR told ____ to take?
8) In the last 12 months, how often did [this condition/the conditions resulting from the (*AIP*)] cause ____ pain or discomfort or upset — all of the time, often, once in a while, or never?
9) When this condition did bother ____, was ____ bothered a great deal, some, or very little?

Note: AIP = accidents, injuries, or poisonings.

APPENDIX B—Categorization Scheme

Condition Category	Checklist Conditions Included
Impairments	
Musculoskeletal impairments	missing limbs, fingers, or toes[a]
	permanent impairment; stiffness or deformity of back or side, limbs, fingers, or toes[a]
Deafness and hearing loss	deafness or trouble hearing in one ear[a]
	deafness or trouble hearing in both ears[a]
Blindness and vision impairment	blindness in one eye[a]
	blindness in both eyes[a]
	crossed eyes[a]
	any other trouble seeing[a]
Speech defects	stammering or stuttering
	any other speech defect
Cerebral palsy	cerebral palsy[a]
Diseases	
Diabetes	diabetes[a]
Sickle cell disease	sickle cell anemia[a]
Anemia	anemia
Asthma	asthma[a]
Eczema and skin allergies	eczema or any other skin allergy[a]
Epilepsy and seizures	epilepsy or convulsions without fever[a]
Arthritis	arthritis or other joint problem[a]
Heart disease	congenital heart disease[a]
	any other heart disease or condition[a]
Frequent ear infections	frequent or repeated ear infections
Frequent diarrhea/bowel trouble	frequent or repeated diarrhea or colitis
	any other persistent bowel trouble
Digestive allergies	any food or digestive allergy
Frequent or severe headaches	frequent or severe headaches, including migraines
Other	mononucleosis
	hepatitis
	meningitis
	rheumatic fever[a]
	seizures associated with fever
	other bone, cartilage, muscle, or tendon problem
	conditions requiring surgery
	conditions lasting more than 3 months

[a]Denotes conditions deemed chronic regardless of date of onset; other conditions deemed chronic if present more than 3 months.

Reductions in HIV Risk–Associated Sexual Behaviors among Black Male Adolescents: Effects of an AIDS Prevention Intervention

John B. Jemmott III, PhD, Loretta Sweet Jemmott, RN, PhD, and Geoffrey T. Fong, PhD

ABSTRACT

Background. The number of reported cases of acquired immune deficiency syndrome (AIDS) is increasing disproportionately among Blacks in the United States. The relatively high incidence of sexually transmitted diseases among Black adolescents suggests the need for AIDS prevention programs to reduce their risk of sexually transmitted human immunodeficiency virus (HIV) infection.

Methods. Black male adolescents (n = 157) were randomly assigned to receive an AIDS risk reduction intervention aimed at increasing AIDS-related knowledge and weakening problematic attitudes toward risky sexual behavior, or to receive a control intervention on career opportunities.

Results. The adolescents who received the AIDS intervention subsequently had greater AIDS knowledge, less favorable attitudes toward risky sexual behavior, and lower intentions to engage in such behavior than did those in the control condition. Follow-up data collected 3 months later revealed that the adolescents who had received the AIDS intervention reported fewer occasions of coitus, fewer coital partners, greater use of condoms, and a lower incidence of heterosexual anal intercourse than did the other adolescents.

Conclusions. These results suggest that interventions that increase knowledge about AIDS and change attitudes toward risky sexual behavior may have salutary effects on Black adolescents' risk of HIV infection. (*Am J Public Health.* 1992;82:372–377)

Introduction

Acquired immune deficiency syndrome (AIDS) has created a major health crisis. In the United States, over 179 000 cases of AIDS had been reported as of May 1991, with over 113 000 deaths attributed to the disease.[1] But these numbers represent only part of the problem. An estimated 1 million people have been infected by the human immunodeficiency virus (HIV), the cause of AIDS.[2,3] These people, too, are likely to develop AIDS. Although HIV can be transmitted perinatally from infected mothers to their newborns, the risk of infection is primarily associated with two kinds of behaviors: (1) the sharing of contaminated hypodermic needles and other drug paraphernalia and (2) the exchange of semen, blood, or vaginal fluids during sexual activities.[1] The most promising method of curtailing the spread of AIDS is to reduce the frequency of high-risk behaviors.

In this paper we report the results of a field experiment employing an intervention designed to reduce the frequency of HIV risk–associated sexual behaviors among inner-city Black male adolescents. Although most reported AIDS cases in the United States involve White homosexual and bisexual men, Blacks are, in fact, disproportionately burdened by AIDS. Only 12% of the nation's population is Black, yet 28% of AIDS patients are Black.[1] Adolescents represent 1% of all reported AIDS cases in the United States. But this statistic may underestimate the potential for HIV infection among adolescents. Adults in their 20s constitute 20% of all reported AIDS cases, and because several years typically elapse between the time a person is infected with HIV and the appearance of clinical signs sufficient to warrant a diagnosis of AIDS, many of these young adults may have acquired the infection as adolescents.[4]

Statistics on rates of sexually transmitted diseases (STDs) and unintended pregnancy suggest that adolescents engage in sexual activities that heighten risk of HIV infection.[5–11] Adolescents have higher rates of syphilis, gonorrhea, and pelvic inflammatory disease than do all other age groups. Inner-city Black adolescents, however, are particularly at risk: STDs are two to three times more common in inner-city populations, and the adolescent pregnancy rate is nearly twice as high among Blacks as among Whites. The prevalence of intravenous drug use in the inner city also heightens the risk of HIV infection for Black adolescent residents. Although the adolescents themselves may not use intravenous drugs—indeed, some data[12] indicate that intravenous drug use among adolescents is rare—they may have sexual relationships with intravenous drug users or with individuals who have had sex with such potentially infected persons.

Despite the potential risk of sexually transmitted HIV infection among inner-city Black adolescents, there is a dearth of intervention studies on behavior change in

John B. Jemmott III is with the Department of Psychology, Princeton University. Loretta Sweet Jemmott is with the College of Nursing at Rutgers, The State University of New Jersey. Geoffrey T. Fong is with the Department of Psychology at the University of Waterloo, Waterloo, Ontario, Canada.

Requests for reprints should be sent to John B. Jemmott III, PhD, Princeton University, Department of Psychology, Green Hall, Princeton, NJ 08544-1010.

This paper was submitted to the Journal February 20, 1991, and accepted with revisions August 22, 1991.

this population. More broadly, research into adolescents' sexual behavior, for the most part, has been conducted on female adolescents.[13] Yet it is also important to understand the dynamics of male adolescents' sexual behavior. Male adolescents are more likely to initiate sexual activities with their partners,[13,14–16] and the male's use of a condom is the best method to reduce the risk of contracting or spreading an STD during sexual activities.

Recent studies provide initial experimental evidence that AIDS prevention interventions emphasizing attitudes, knowledge, and skill-building can increase condom procurement among Black male adult STD patients[17] and self-reported condom use among White male homosexuals.[18] In an AIDS intervention study[19] of Black and Hispanic adolescents, those in the intervention conditions, compared with those not receiving an intervention, showed increased ratings of the value of AIDS education and reduced approval of intravenous drug use 1 month after the intervention. DiClemente and colleagues[20] tested an AIDS education intervention on classes within middle and high schools in San Francisco and found that, compared with control classes, those that received the AIDS instruction had greater AIDS knowledge and less fear of being infected by classmates who have the disease. DiClemente et al.[20] urged behavioral assessments in future intervention studies on adolescents.

In the present study we examined whether Black male adolescents who were randomly assigned to an AIDS risk reduction intervention would have greater knowledge and less positive attitudes and intentions regarding risky sexual behavior than would those assigned to a control intervention on career opportunities. Most important, we also examined whether risky sexual behaviors would decline in the intervention condition, relative to the control condition, during the 3 months after the intervention.

Methods

Study Sample

The participants were 157 Black male adolescents (mean age = 14.64 years, SD = 1.66) from Philadelphia, Pa, who were recruited from among outpatients at a medical clinic in West Philadelphia (44%); students attending the 10th, 11th, and 12th grade assemblies at a local high school (32%); and adolescents at a local YMCA (24%). (The results do not differ as a function how subjects were recruited.) About 97% of the participants were currently enrolled in school. The mean number of years of education of their mothers, as reported by participants, was 13.82 (SD = 3.17). Few participants (4.5%) reported ever sharing needles or ever having receptive anal intercourse (2.3%), or having sexual relationships with males exclusively (1.6%) or with both males and females (0.8%). Their chief HIV risk was from heterosexual activities. About 33.9% reported having more than one coital partner in the previous 3 months, and 12.8% of the respondents indicated that they had had heterosexual anal intercourse during that period. About 20.9% of the respondents who had had coitus in the previous 3 months reported that they never used condoms during those experiences, and only 30.2% reported always using condoms. The adolescents were offered $40 for participating: $15 for taking part in the intervention and $25 for attending the 3-month follow-up. Of the original participants, 150 completed the 3-month follow-up questionnaire, for a return rate of 96%. The return rate was not different between experimental conditions.

Procedure

The study was approved by the human subjects committee of Princeton University. Adolescents and parents signed separate consent forms. They were given descriptions of the study that took into account the fact that the adolescents would be randomly assigned to an AIDS risk reduction intervention or to a career-opportunities intervention. The study was described as a risk reduction project, designed "to understand Black male youths' behaviors that may create risks such as unemployment, truancy, teenage pregnancy, and sexually transmitted diseases, especially AIDS, and to find ways to teach Black male youth how to reduce these risks."

The intervention session was held at a local school on Saturday, October 15, 1988. That morning, the participants completed preintervention questionnaires, which took about 90 minutes. Immediately before completing the questionnaires, the participants signed an agreement indicating that they understood that it was important to answer the questions carefully and honestly, that their answers would be confidential, and that their names would not be put on their questionnaires. The participants also signed such agreements before completing postintervention and follow-up questionnaires. Research indicates that this procedure increases the accuracy of self-reports. While the participants completed the preintervention questionnaires, they were stratified by age and randomly assigned within age to either the AIDS condition or the career-opportunities condition and to one of 27 small groups (14 in the AIDS condition and 13 in the control condition) led by a trained male or female facilitator. A total of 85 adolescents were assigned to the AIDS condition, and 72 were assigned to the control condition. The sample sizes are smaller in some analyses because of attrition or because subjects gave no response to the question.

AIDS risk-reduction condition. Participants in the AIDS risk reduction condition received a 5-hour intervention designed to increase their knowledge of AIDS and STDs and to weaken problematic attitudes toward risky sexual behaviors. The intervention included information about risks associated with intravenous drug use and specific sexual activities. Videotapes, games, exercises, and other culturally and developmentally appropriate materials were used to reinforce learning and to encourage active participation. All materials had been extensively pilot tested and were selected not only to provide accurate information, but to do so in ways that would be interesting to inner-city Black male adolescents. For example, one video, "The Subject is AIDS," was narrated by a Black woman and had a multiethnic cast. In one game, "AIDS Basketball,"[21] participants in small-group sessions were divided into teams that earned points for correctly answering factual questions on AIDS. In one exercise, "Uncle Bill's Advice Column,"[21] pairs of adolescents wrote a response to a letter to Uncle Bill about AIDS or risky behavior and then read it to the rest of the group for discussion. A condom exercise focused on the correct use of condoms. The participants also engaged in role-playing situations depicting potential problems in trying to implement safer sex practices, including abstinence. A more detailed description of the intervention is available from the authors.

Career-opportunities condition. An important concern in designing this experiment was to control for Hawthorne effects to reduce the likelihood that effects of the AIDS intervention could be attributed to nonspecific features, including group interaction and special attention. A waiting-list or no-treatment control condition would not control for these confounding variables and consequently was seen

as undesirable. In addition, for ethical and methodological reasons, every effort was made to ensure that the experiences of control participants were as valuable and enjoyable as those of experimental participants.[22] Accordingly, participants in the control condition received an intervention concerning career planning and opportunities. Structurally similar to the AIDS intervention, it also lasted 5 hours and used videotapes, small-group discussions, exercises, and games. The participants viewed the film "Career and You." In one game, "The Career Game," participants in small groups were divided into teams that earned points for correctly answering questions about careers. They participated in a value-awareness exercise, and they did a career analysis.

After the interventions, all participants completed the postintervention questionnaires. This took about an hour. To reduce the effects of demand characteristics, the questionnaires were administered by research assistants, not by the facilitators.

Facilitators

The facilitators were 27 Black adults (11 women and 16 men) with a mean age of 36.2 years (SD = 6.8). The AIDS facilitators had backgrounds in human sexuality education, mental health counseling, AIDS education, nursing, or social work. The career-opportunities facilitators had backgrounds in education, career guidance, or business. All facilitators had at least a 4-year college degree. A majority had experience with small-group facilitation.

One week before the intervention, the facilitators received 6 hours of training. AIDS and careers facilitators were trained at the same time, but in separate groups. By discussing potential problems and issues as a group, the facilitators were able to fashion common responses. The training emphasized that all facilitators of the same intervention must implement it in the same way so as to reduce facilitator effects.

Measures

Risky sexual behavior during the previous 3 months. On the preintervention and follow-up questionnaires, the participants answered a series of questions about their sexual practices in the previous 3 months, including coitus, multiple sexual partners, condom use, and heterosexual anal intercourse. At each assessment, participants' responses to the individual items were converted to z scores and then averaged to form a scale on which higher scores indicated higher incidence of risky sexual behavior. Coefficient α for the scale was .67 at preintervention and .72 at follow-up.

Intentions and attitudes regarding risky sexual behavior in the next 3 months. The theory of reasoned action and a growing empirical literature suggest that attitudes and intentions are related to behavior, including HIV risk–associated sexual behavior.[23–27] The participants' attitudes and intentions regarding engaging in the aforementioned sexual practices in the next 3 months were measured at preintervention, immediately postintervention, and at the 3-month follow-up. To measure intentions, participants were asked to rate the likelihood that they would engage in each behavior in the next 3 months on a scale from "extremely unlikely" (1) to "extremely likely" (7). At each time point, the ratings were averaged to form scales on which higher numbers indicated stronger intentions to engage in risky sexual behavior in the next 3 months. Coefficient α was .68 at preintervention, .72 at postintervention, and .70 at follow-up. To measure attitudes, participants were asked to rate their attitudes toward engaging in each behavior in the next 3 months on a scale from "extremely negative" (1) to "extremely positive" (7). At each time point, the ratings were averaged to form scales on which higher numbers indicated more favorable attitudes toward engaging in risky sexual behavior in the next 3 months. Coefficient α was .63 at preintervention, .68 at postintervention, and .63 at follow-up.

AIDS and STD knowledge. Participants answered factual true-false questions regarding AIDS and STDs. Thirty items concerned AIDS and were taken from scales previously used in research on adolescents by DiClemente et al.[20,28] Thirty additional items concerned information about either AIDS or STDs in general that is not tapped by the DiClemente measures. To ensure the content validity of the scale, six AIDS experts—scientists and clinicians involved in AIDS-related research, patient care, or education—rated the information in each item on a 3-point scale from "unimportant" (0) to "important" (1) to "essential" (2) for laypersons to know. Eighteen items were seen as essential (mean ratings = 2.0) by every judge, and 39 items had mean ratings in the important to essential range (mean ratings $\geq$ 1.00). Three items the judges considered either ambiguous or unimportant were deleted. The number correct out of 57 was calculated from the preintervention, postintervention, and follow-up questionnaires. Coefficient α was .73, .89, and .82, respectively.

Other measures. On the postintervention questionnaire, the participants answered a series of questions about their perceptions of the intervention activities. The facilitators completed a questionnaire regarding their impressions of the experience of the adolescents in their small groups and their success at implementing the intervention. The Marlowe-Crowne Social Desirability Scale[29] was used to measure the tendency of participants to describe themselves in favorable, socially desirable terms.

Results

Intervention Integrity

Before considering effects of the AIDS intervention on knowledge, attitudes, intentions, and behavior, it is important to establish that the participants in the two conditions were equally involved in their interventions and found them to be valuable and enjoyable. There were no differences between conditions in the participants' ratings of their liking for the intervention, their learning from the activities, whether the intervention would help them in the future, their general emotional reactions toward the intervention, or their participation level. The only significant difference between the groups was that those in the AIDS condition would be more likely to recommend the project to other teenagers than would those in the careers condition—mean ratings of 4.18 vs 3.78 on a 5-point scale, F (1, 152) = 4.43, $P < .04$. The facilitators' ratings of their small groups were consistent with the participants' self-ratings: They revealed a high degree of attentiveness and interest, with no differences between conditions. In addition, there was evidence that the facilitators had been successful in presenting the interventions in the same way. Analyses on knowledge, attitudes, intentions, and behavior revealed no differences that could be attributed to the particular small group in which the adolescents participated, independent of experimental condition.

Knowledge, Attitudes, and Intentions

The data were analyzed with Condition × Gender of Facilitator analyses of covariance,[30] controlling for preintervention measures of the relevant crite-

rion. Table 1 contains the adjusted postintervention means. Immediately after the intervention, participants in the AIDS condition had greater knowledge about AIDS, F (1, 151) = 19.58, $P < .0001$, expressed less favorable attitudes toward risky sexual behaviors, F (1, 150) = 8.42, $P < .004$, and reported weaker intentions to engage in such behavior than did their counterparts in the career-opportunities condition, F (1, 150) = 17.45, $P < .0001$. The Condition × Gender of Facilitator interaction on knowledge was significant, F (1, 151) = 4.26, $P < .04$, indicating that the degree of increase in knowledge caused by the AIDS intervention was greater with male facilitators than with female facilitators.

Three-month follow-up. Three months after the intervention, participants in the AIDS condition still scored higher on AIDS knowledge, F (1, 147) = 9.46, $P < .003$, and reported weaker intentions to engage in risky sexual behavior in the next 3 months, F (1, 144) = 7.58, $P < .007$, than did those in the career-opportunities condition. Table 2 contains the adjusted means. There was a nonsignificant trend toward less favorable attitudes regarding risky sexual behavior among those in the AIDS condition than among those in the control condition, F (1, 144) = 2.82, $P < .10$. However, a significant Condition × Gender of Facilitator interaction, F (1, 144) = 8.91, $P < .003$, indicated that the effect of the AIDS intervention on lowering attitudes was greater with female facilitators than with male facilitators.

Risky Sexual Behavior

As Table 2 shows, when preintervention reports of risky sexual behavior are controlled for, participants in the AIDS condition reported engaging in less risky sexual behavior in the 3 months following the intervention than did participants in the career-opportunities condition, F (1, 135) = 6.48, $P < .01$. The Condition × Gender of Facilitator interaction was significant, F (1, 135) = 4.19, $P < .04$, indicating that the reduction in risky sexual behavior caused by the AIDS intervention was greater with female facilitators than with male facilitators.

Table 3 shows that the effects of the intervention were fairly consistent across different sexual behaviors. Although adolescents in the AIDS condition were not significantly ($P < .10$) more likely than those in the career-opportunities condition to practice complete abstinence in the 3 months postintervention, they did report having coitus on fewer days, with fewer women, and with fewer women who were involved in sexual relationships with other men. Subjects in the AIDS condition also reported fewer occasions of coitus without a condom and were less likely to report having anal intercourse with a woman in the 3 months postintervention than were the other adolescents. Condition × Gender of Facilitator interactions of the same form as on the aggregate risky behavior index were statistically significant on number of days on which the participants engaged in coitus, number of days on which the participants had coitus without using a condom, and experience with heterosexual anal intercourse in the past 3 months.

TABLE 1—Adjusted[a] Postintervention Means, by Experimental Condition

Variable	AIDS Prevention Condition	Control Condition	Difference	95% CI of Difference
Knowledge	46.22 (85)	40.37 (71)	5.85	3.23, 8.48
Attitudes	3.02 (83)	3.48 (72)	−.46	−.77, −.15
Intentions	2.83 (83)	3.40 (72)	−.57	−.84, −.30

Note. Numbers of respondents are shown in parentheses. CI = confidence interval. Higher numbers indicate greater AIDS knowledge, more positive attitudes toward behaviors that increase risk of sexually transmitted HIV infection, and stronger intentions to engage in such behaviors.
[a] For each variable, the preintervention measure is partialled out of the postintervention measure.

TABLE 2—Adjusted[a] 3-Month Follow-up Means by Experimental Condition

Variable	AIDS Prevention Condition	Control Condition	Difference	95% CI of Difference
Knowledge	47.20 (83)	44.40 (67)	2.8	.72, 4.88
Attitudes	3.13 (82)	3.38 (67)	−.25	−.54, .04
Intentions	2.87 (82)	3.30 (67)	−.43	−.74, −.12
Risky behavior	−.12 (76)	.24 (62)	−.36	−.64, −.08

Note. Numbers of respondents are shown in parentheses. CI = confidence interval. Higher numbers indicate greater AIDS knowledge, more positive attitudes toward behaviors that increase risk of sexually transmitted HIV infection, stronger intentions to engage in such risky behaviors, and more reports of risky sexual behavior.
[a] For each variable, the preintervention measure is partialled out of the 3-month follow-up measure.

Social Desirability Bias

Marlowe-Crowne Social Desirability Scale scores were unrelated to preintervention self-reports, 3-month follow-up self-reports, or amount of change in self-reports of risky sexual behavior. In addition, the scores were unrelated to preintervention, postintervention, or follow-up intentions and attitudes or to changes in these variables.

Discussion

Given the widely recognized potential risk of sexually transmitted HIV among inner-city Black adolescents, the results of the present experiment are encouraging. They suggest that interventions that increase knowledge about AIDS and change attitudes toward risky sexual behavior may have salutary effects on Black adolescents' risk of HIV infection. Compared with their counterparts in the control condition, the inner-city Black male adolescents who were randomly assigned to an AIDS intervention subsequently had lower intentions of engaging in sexual behaviors that would increase their risk of contracting or spreading HIV infection. These adolescents also had less favorable attitudes toward risky sexual behaviors and greater AIDS knowledge. Analyses on data from the 3-month follow-up revealed that the adolescents who received the AIDS intervention also reported engaging in less risky sexual behavior following the intervention than did those who received the control intervention. Moreover, the effects of the intervention on intentions regarding risky sex-

TABLE 3—Adjusted[a] Means for Specific Sexual Behaviors, by Experimental Condition

Variable	AIDS Prevention Condition	Control Condition	Difference	95% CI of Difference
Coitus[b]	.48 (62)	.60 (53)	−.12	−.27, 0.3
No. days respondent had coitus	2.15 (53)	5.48 (47)	−3.32	−5.78, −.89
No. coital partners	.85 (63)	1.79 (52)	−.93	−1.53, −.33
No. coital partners involved with other men	.19 (60)	1.75 (48)	−1.55	−2.67, −.43
Rated frequency of condom use[c]	4.35 (22)	3.50 (21)	.85	.14, 1.56
No. days respondent did not use condom during coitus	.64 (58)	2.38 (48)	−1.73	−2.86, −.60
Heterosexual anal sex[b]	.07 (68)	.27 (49)	−.19	−.32, −.06
No. days respondent had heterosexual anal sex	.36 (64)	.92 (46)	−.55	−1.17, .07
No. female anal sex partners	.13 (65)	.61 (49)	−.47	−.86, .08

Note. Numbers of respondents are shown in parentheses. CI = confidence interval.
[a] For each variable, the preintervention measure is partialled out of the 3-month follow-up measure.
[b] Coitus and heterosexual anal sex are dummy coded so that 1 = "yes" and 0 = "no."
[c] Frequency of condom use was rated on a 5-point Likert-type scale from 1 = "never" to 5 = "always."

ual behavior and AIDS-related knowledge were sustained over the 3 months. The effects on attitudes observed immediately postintervention were evident at the follow-up among participants who had a female facilitator.

These results cannot be explained as a simple result of the special attention received by the adolescents in the AIDS intervention condition. The two interventions were matched in length and involved similar kinds of activities. Moreover, the participants' evaluations of their experiences were quite similar across the interventions. Although career-opportunities subjects did not learn about AIDS, they did gain knowledge that would be valuable to them in the future, given the high unemployment among inner-city Black adolescents.

By its very nature, risky sexual behavior is private behavior and consequently must be assessed using self-report measures, which may be biased.[31] We employed several techniques to make it less likely that the subjects would minimize or exaggerate reports of their sexual experiences: (1) code numbers rather than names were used on the questionnaires, (2) facilitators were not involved in any way in the administration of questionnaires, (3) the importance of responding honestly was emphasized, and (4) participants were assured that their responses would be kept confidential. Furthermore, if concern about how they would be viewed by others influenced respondents' reports of their sexual behavior, those whose need for social approval was stronger[29] might have differed from the other adolescents in self-reported risky sexual behavior or in the change in their reports after the AIDS intervention. Responses, however, were unrelated to social desirability bias. Nevertheless, because it is difficult to validate self-reported sexual behavior,[31] interpretations of the present findings should take into consideration the possibility that the participants' self-reports might have been inaccurate.

It might be argued that the gender of facilitators and group members should be matched to enhance an intervention's effectiveness. Consonant with this argument, the AIDS intervention caused greater increases in postintervention AIDS knowledge when the facilitator was male than when the facilitator was female. But this advantage of male facilitators was not evident on postintervention attitudes or intentions, and it vanished at follow-up. In fact, 3-month follow-up data revealed that the AIDS intervention caused less self-reported risky behavior and less positive attitudes toward such behavior when the facilitator was female than when the facilitator was male. These results provide scant support for the view that matching the gender of facilitator and intervention recipient enhances the effectiveness of AIDS interventions with Black male adolescents.

One common argument against AIDS education programs for adolescents and children has been that exposing them to information about sex will encourage them to engage in sexual activity. Our data, however, provide some evidence that the opposite may be true. Adolescents who received the AIDS intervention were *less* likely to engage in sexual activity, and those who did were more likely to engage in safer sexual activity. Thus, the common fear that providing adolescents with information about AIDS will result in greater sexual activity is perhaps simply a fear.

The present study focused on sexual behaviors that heighten risk of HIV infection. However, the overwhelming majority of the subjects did not report any homosexual or bisexual experiences. Sharing needles and other intravenous drug use paraphernalia would also increase risk of HIV infection, but few of these subjects reported engaging in such behavior. The important question of whether interventions of the type used in the present study would be effective with Black male adolescents who engage in homosexual or bisexual activities or who use intravenous drugs will have to be addressed in future studies. More broadly, research on strategies to change HIV risk–associated behavior in diverse populations is still in its infancy. We are optimistic, however, that by conducting research along these lines, it may be possible to curb the spread of AIDS. □

Acknowledgments

This research was supported in part by American Foundation for AIDS research grants 000667 and 000952 and National Institute of Child Health and Human Development grant HD24921.

Some of these data were presented at the meeting of the Society of Behavioral Medicine, April 1990, Chicago, Ill.

The authors gratefully acknowledge the contributions to this research project of Leonard Johnson, Konstance McCaffree, and Dionne Jane Latham, and the helpful suggestions of Paula Braverman, David Brody, Ann Marie Carosella, Robert T. Croyle, James M. Jones, Pamela Kato, Caryn Lerman, David Schor, Melvin Snyder, and Mark Zanna regarding an earlier version of this article.

References

1. Centers for Disease Control. HIV/AIDS Surveillance Report, June, 1–18. Atlanta, Ga: Center for Infectious Diseases, Centers for Disease Control; 1991.
2. Centers for Disease Control. Human immunodeficiency virus in the United States: a review of current knowledge. *MMWR*. 1987;36(Suppl. 5–6), 1–48.
3. Centers for Disease Control. CDC HIV/AIDS Prevention Newsletter, February. Atlanta, Ga: Centers for Disease Control; 1991.
4. Curran JW, Jaffe HW, Hardy AM, Morgan W, Selik R, Dondero T. Epidemiology of HIV infection and AIDS in the United States. *Science*. 1988;239:610–616.
5. Bell TA, Holmes KK. Age specific risks of syphyllis, gonnorrhea, and hospitalized pelvic inflammatory disease in sexually experienced U.S. women. *Sex Transm Dis*. 1984;7:291.
6. Hatcher RA, Stewart FH, Trussell J, et al. *Contraceptive Technology 1990–1992*. New York, NY: Irvington; 1990.
7. Hayes CD, ed. *Risking the Future: Adolescent Sexuality, Pregnancy, and Childbearing*, Volume 1. Washington, DC: National Academy Press; 1987.
8. Hofferth SL, Hayes CD, eds. *Risking the Future: Adolescent Sexuality, Pregnancy, and Childbearing*, Volume 2. Washington, DC: National Academy Press; 1987.
9. Jones EF, Forrest JD, Goldman N, et al. *Teenage Pregnancy in Industrialized Countries*. New Haven, Conn: Yale University Press; 1986.
10. Pratt W, Mosher W, Bachrach C, Horn M. Understanding U.S. fertility: findings from the National Survey of Family Growth, Cycle III. *Popul Bull*. 1984;39:1–42.
11. Zelnik M, Kantner JF. Sexual activity, contraceptive use and pregnancy among metropolitan-area teenagers: 1971–1979. *Fam Plann Perspect*. 1980;12:230–238.
12. Turner CF, Miller HG, Moses LE, eds. *AIDS, Sexual Behavior and Intravenous Drug Use*. Washington, DC: National Academy Press; 1989.
13. Weddle KD, McKenry PC, Leigh GK. Adolescent sexual behavior: trends and issues in research. *J Adol Res*. 1988;3:245–257.
14. Chilman CS. *Adolescent Sexuality in a Changing American Society: Social and Psychological Perspectives*. Bethesda, Md: Department of Health, Education, and Welfare Publication No. NIH 79-1426; 1979.
15. Freeman EW, Rickels K, Huggins GR, Mudd EH, Garcia C, Dickens HO. Adolescent contraceptive use: comparisons of male and female attitudes and information. *Am J Public Health*. 1980;70:790–797.
16. Koch PB. The relationship of first intercourse to later sexual functioning concerns of adolescents. *J Adol Res*. 1988;3:345–362.
17. Solomon M, DeJong W. Preventing AIDS and other STDs through condom promotion: a patient education intervention. *Am J Public Health*. 1989;79:453–458.
18. Kelly J, Lawrence J, Hood H, Brasfield T. Behavioral intervention to reduce AIDS risk activities. *J Consult Clin Psychol*. 1989;57:60–67.
19. Schinke SP, Gordon AN, Weston RE. Self-instruction to prevent HIV infection among African-American and Hispanic-American adolescents. *J Consult Clin Psychol*. 1990;58:432–436.
20. DiClemente RJ, Pies CA, Stoller EJ, et al. Evaluation of school-based AIDS education curricula in San Francisco. *J Sex Res*. 1989;26:188–198.
21. Yarber W. AIDS: What young adults should know. Bloomington, Ind: Indiana State Board of Health; 1987.
22. Cook TD, Campbell DT. *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago, Ill: Rand McNally; 1979.
23. Ajzen I, Fishbein M. *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice-Hall; 1980.
24. Fishbein M, Middlestadt SE. Using the theory of reasoned action as a framework for understanding and changing AIDS-related behaviors. In: Mays V, Albee G, Schneider S, eds. *Primary Prevention of AIDS: Psychological Approaches*. Beverly Hills, Calif: Sage; 1989:93–110.
25. Jemmott LS, Jemmott JB III. Sexual knowledge, attitudes, and risky sexual behavior among inner-city Black male adolescents. *J Adol Res*. 1990;5:346–369.
26. Jemmott LS, Jemmott JB III. Applying the theory of reasoned action to AIDS risk behavior: condom use among Black women. *Nurs Res*. 1991;40:228–234.
27. Jorgensen SR, Sonstegard JS. Predicting adolescent sexual and contraceptive behavior: an application and test of the Fishbein model. *J Marriage Family*. 1984;46:43–55.
28. DiClemente RJ, Zorn J, Temoshok L. Adolescents and AIDS: a survey of knowledge, beliefs, and attitudes about AIDS in San Francisco. *Am J Public Health*. 1988;76:1443–1445.
29. Crowne D, Marlowe D. *The Approval Motive*. New York, NY: John Wiley & Sons; 1964.
30. Huitema, BE. *The Analysis of Covariance and Alternatives*. New York, NY: John Wiley & Sons; 1980.
31. Catania J, Gibson DR, Chitwood DD, Coates TJ. Methodological problems in AIDS behavioral research: influences on measurement error and participation bias in studies of sexual behavior. *Psychol Bull*. 1990;108:339–362.

Birth Weight and Perinatal Mortality: The Effect of Gestational Age

ABSTRACT

Background. The strong association between birth weight and perinatal mortality is due both to gestational age and to factors unrelated to gestational age. Conventional analysis obscures these separate contributions to perinatal mortality, and overemphasizes the role of birth weight. An alternative approach is used here to separate gestational age from other factors.

Methods. Data are from 400 000 singleton births in the Norwegian Medical Birth Registry. The method of Wilcox and Russell is used to distinguish the contributions to perinatal mortality made by gestational age and by relative birth weight at each gestational age.

Results. Gestational age is a powerful predictor of birth weight and perinatal survival. After these effects of gestational age are controlled for, relative birth weight retains a strong association with survival.

Conclusions. Current public health policies in the United States emphasize the prevention of low birth weight. The present analysis suggests that the prevention of early delivery would benefit babies of all birth weights. (*Am J Public Health.* 1992;82:378–382)

Allen J. Wilcox, MD, PhD, and Rolv Skjærven, PhD

Introduction

Birth weight is the single strongest predictor of infant survival. However, its position in the causal pathway is unclear. One determinant of birth weight is gestational age: as the fetus matures, it grows. Some of the association of weight with survival presumably reflects this maturation. There is additional variability in weight that is unexplained by gestational age. This second type of birth weight variability is also strongly related to infant survival, but by biological mechanisms less well understood. The purpose of this paper is to separate mortality as related to gestational age from mortality as related to relative birth weight within fixed gestational age. In the process of such a separation, factors related to gestational age can be separated from other factors that contribute to perinatal mortality. This produces a picture of mortality that is different from the traditional weight-versus-age multivariate approach, a picture which may better reflect the underlying causal pathways.

Methods

Data are from the Medical Birth Registry of Norway, which includes all live births and fetal deaths after 16 weeks of gestation registered in Norway since 1967. We selected all singleton first births through 1984 with gestational ages of 28 or more weeks, totaling 413 051. Birth weights were grouped into 100-g categories, and gestational ages were grouped by completed weeks since last menstrual period. Birth weight data were missing for 0.2% of births and gestational age was missing for an additional 4.3%. Those births were excluded, leaving 394 386 births (Table 1). We defined perinatal mortality as fetal deaths plus deaths in the first week of life. Total perinatal mortality in this study group was 12.9 per thousand.

Our method of analysis adjusts for birth weight with the modified form of direct standardization suggested by Wilcox and Russell.[1,2] The modification is first to adjust all birth weights to a *z*-score scale. (A *z* score expresses birth weight in standard-deviation units relative to mean weight.) This *z* scale is based on the underlying Gaussian distribution of births in each group being compared. Standardization is carried out on the *z* scale of birth weight rather than on a scale of absolute birth weight. Such an approach allows the comparison of mortality for babies of the same *relative* weight (that is, relative to the Gaussian distribution of weight for the particular group), rather than for babies of the same absolute weight. When applied to an analysis of gestational age, this method identifies babies that are relatively small or large for their gestational age and compares them with babies of the same relative size at other gestational ages.

We estimate the *z* scale by fitting a Gaussian curve to the birth weight distri-

Allen J. Wilcox is with the Epidemiology Branch of the National Institute of Environmental Health Sciences, Research Triangle Park, NC. Rolv Skjærven is with the Section for Medical Informatics and Statistics, University of Bergen, Bergen, Norway.

Requests for reprints should be sent to Allen J. Wilcox, MD, PhD, Epidemiology Branch, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709, or to Rolv Skjærven, PhD, Section for Medical Informatics and Statistics, University of Bergen, Bergen, Norway.

This paper was submitted to the journal May 9, 1991, and accepted with revisions September 16, 1991.

Editor's Note. See related Editorial by Kiely and Susser on page 343 of this issue.

bution within each gestational age stratum. Figure 1 shows the distribution of birth weights in each stratum. Birth weight distributions between 38 and 46 weeks are almost exactly Gaussian. Below 38 weeks, two systematic deviations from the Gaussian emerge. First, as gestational age decreases there are increasing numbers of small babies in the lower tail of the distribution, outside the Gaussian. Many of these are presumably pathological conceptuses, including macerated stillbirths who stopped growing some time before delivery. (At 28 to 31 weeks, the distribution in the lower tail is obscured by the grouping of births less than 1000 g.)

Second, there is an increasing number of relatively heavy births as gestational age decreases. This pattern has been frequently observed and is attributed to errors in measurement of gestational age.[3,4] Gestational age defined by last menstrual period can be in error for at least two physiologic reasons: the estimated gestation length can be too short when bleeding in early pregnancy is mistaken for the last menstrual period, or too long when conception follows an extended follicular phase. The excess of heavier babies at early gestational ages is thought to represent babies whose recorded age is shorter than their true gestational age, and who therefore actually belong in one of the heavier distributions occurring at later gestations. There is presumably some misclassification of gestational age in every stratum, but the unequal distribution of births among the strata makes misclassification more apparent in some strata than in others. A small rate of misclassification among the large numbers of births around term would produce a considerable excess of heavy babies among the small numbers born into the preterm groups.

Deviations from the Gaussian distribution cause some difficulties in estimating parameters of the underlying Gaussian. A FORTRAN program has been developed to estimate these parameters in the presence of excess small births.[5,6] (This program is available upon request.) The FORTRAN procedure was suitable for most gestational-age strata, but did not work well at gestations of less than 38 weeks owing to the excess of heavy births. In those cases we have fit a Gaussian curve more approximately to the modal region of each distribution and extrapolated a standard deviation from later strata (Figure 1). Parameters of the Gaussian distribution for each gestational age stratum are shown in Table 2.

TABLE 1—Distribution of Births and Deaths by Gestational Age, with Gestational-age-Specific Relative Risks for Perinatal Death (Relative to Births at 40 weeks) before and after Adjustment for Birth weight

Gestational Age (weeks)	All Births	Perinatal Deaths	Perinatal Mortality[a]	Relative Risk[b]	Birth-weight-adjusted Relative Risk
28–31	2834	1034	364.9	85	91
32–33	3231	612	189.4	44	45
34	2924	318	108.8	25	21
35	4692	283	60.3	14	12
36	7824	285	36.4	8.5	6.4
37	14 172	332	23.4	5.4	4.2
38	30 254	370	12.2	2.8	2.3
39	67 970	411	6.0	1.4	1.3
40	104 126	450	4.3	1.0	1.0
41	88 946	467	5.3	1.2	1.2
42	46 457	301	6.5	1.5	1.6
43	12 810	128	10.0	2.3	2.4
44	4692	59	12.6	2.9	3.0
45	2332	29	12.4	2.9	2.6
46+	1122	10	8.9	2.1	2.1
Total	394 386	5089	12.9		

[a]Per 1 000 births.
[b]Relative to mortality at 40 weeks.

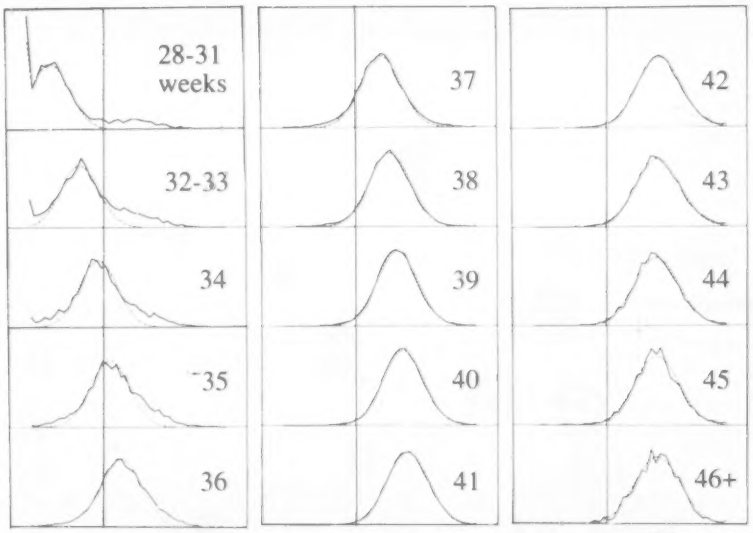


FIGURE 1—Frequency distributions of birth weight stratified by gestational age (solid curves), with underlying Gaussian distributions (dashed curves). A line at 2500 g is provided for reference.

Once parameters for each Gaussian were in hand, the mean and standard deviation for each gestational-age group were used to rescale birth weight to a z score. Weight-specific mortality rates for each group were also adjusted to the z scale in preparation for direct standardization. The choice of a distribution for the standard is not crucial. In general, the standard distribution should be similar to the actual distributions in the groups being analyzed. We chose a Gaussian distribution as our standard because it meets this criterion (Figure 1), but other distributions produce similar results. This standard distribution was applied to the weight-specific mortality rates for each gestational-age stratum. The risk of mortality for each gestational age was then expressed as a relative risk, with babies born at 40 weeks as the reference group.

TABLE 2—Gestation-specific Parameters of the Gaussian Birth Weight Distribution

Gestational Age weeks	Parameters of the Underlying Gaussian	
	Mean, g	SD, g
28–31	1425	400
32–33	2010	400
34	2375	400
35	2610	415
36	2845	430
37	3050	430
38	3225	435
39	3380	430
40	3500	430
41	3595	440
42	3645	460
43	3605	470
44	3560	470
45	3560	465
46+	3585	480

Results

A strong gradient of risk was observed across gestational age after birth weights were adjusted to a *z* scale. Babies at the earliest stratum had a relative risk of 91 (Table 1). Relative risks calculated by this method are similar to those for gestational age unadjusted for birth weight.

The same results are presented graphically in Figures 2 and 3. First, weight-specific mortality rates for nine selected gestational-age strata are shown before adjustment to a *z* scale. (Mortality rates have been smoothed by grouping weight into 300-g categories.) These curves have the usual features of weight-specific mortality: rates are highest at the lowest birth weights, fall rapidly as birth weights increase, and then rise slightly at the highest birth weights.[7]

At lower weights, mortality rates are similar across the various gestational age groups, whereas at higher weights the rates tend to diverge. This pattern changes with adjustment of the birth weight scale. Figure 3 shows the identical mortality curves after birth weight has been converted to the *z* scale. Curves that formerly converged at the lower weights are now separate and roughly parallel over the whole range of birth weights. The distance between any two curves in Figure 3 represents the log of the relative risk associated with the corresponding difference in gestational age. The adjusted relative risks in Table 1 summarize the differences between the lowest mortality curve (at 40 weeks) and each of the other mortality curves in Figure 3.

Figure 4 inverts the display to show perinatal mortality risk by gestational age for babies at given *relative* birth weights. Thus, the top curve shows mortality rates for that high-risk group of babies who are four standard deviations below the mean weight for their gestational age. The next-to-bottom curve shows the mortality experience of babies who are at the mean weight for their gestational age. Within each birth weight group there is a strong gradient of mortality risk with gestational age, with the lowest risk occurring at 40 weeks.

Discussion

Growth is a natural part of the gestational maturation of the fetus. When we refer here to "gestational age" we mean not merely age itself but the tightly correlated phenomena of growth and maturation. Gestational age is a major contributor to birth weight, but it is only one of several contributors. There is another set of factors, less well understood, that produces variation in birth weight within each stratum of gestational age. Some authors attribute this variation entirely to the quality of fetal growth. We regard this interpretation as potentially misleading[8] and prefer the more neutral description "relative weight at each gestational age." This expression is a generalized form of the clinical terms "small for gestational age" or "large for gestational age."

The usual analysis of birth weight and gestational age finds that most of the association with mortality is through birth weight. This was well illustrated by Susser, Marolla, and Fleiss, who reported that when gestational age and weight are analyzed simultaneously, birth weight accounts for 90% of the variance of perinatal mortality, whereas gestational age accounts for barely 5%.[9]

Such conclusions are based on data in the form shown here in Figure 2. Arranged in this way, the mortality curves at different gestational ages are nearly indistinguishable across the lowest birth weights (where rates are highest). For example, babies at 2250 g have a similar risk of death whether their gestational age is 35 or 37 or 40 weeks (Figure 2). Although this analysis may seem to isolate the separate effects of gestational age and birth weight, it does not uncover the causal pathway. Susser and his colleagues were careful not to draw such conclusions, but these data have suggested to some that birth weight is fundamentally more important to survival than gestational age.[10]

If we compare babies of the same *relative* weight at each gestational age, a different picture emerges. Figure 3 shows mortality rates for babies at various relative weights (expressed as standard deviations from the mean at each gestational age). In effect, this compares all small-for-gestational-age babies (or all average-weight babies, or all large-for-gestational-age babies) across different gestational ages. Actual comparison of small-for-gestational-age mortality risks across gestational-age strata shows the same strong effect of gestational age on mortality.[11]

Babies of 2250 g have similar mortality risk regardless of gestational age. But these babies differ considerably in their relative size (Figure 3). At 35 weeks, a baby of 2250 g is about one standard de-

FIGURE 2—Weight-specific mortality rates for nine representative strata of gestational age.

viation below the Gaussian mean weight. At 37 weeks, a baby of that weight is about two standard deviations below the mean, and at 40 weeks it is about three standard deviations below the mean. (See circled points on Figure 3). The similar mortality risk of these three babies is the result of two opposing trends: relative size is worse, and gestational age is better.

The usual analytic methods, in which 2250-g babies are compared at different gestational ages, do not take into account the full benefit that would accrue to a preterm baby from extended gestation. Any strategy of prevention or intervention implicitly assumes that extended gestation is accompanied by fetal growth, that is, that a fetus' *relative* weight stays roughly the same as gestation advances. To measure the advantage of longer gestation, the survival of a 2250-g baby at 35 weeks should be compared with the survival of a baby who has grown at the expected rate in the interim. Such a comparison is accomplished by the adjustment to a relative birth weight scale. In Figures 3 and 4 the effects of relative size and gestational age are separated, showing the relationship of each to survival.

The parameters estimated for the Gaussian distributions of weight at the earlier gestational ages are subject to judgment. Even so, the uncertainty regarding parameters used for z scales is not crucial to the conclusion. As an illustration, births at 35 weeks had a relative risk of 12, compared with births at 40 weeks. If we vary the mean of the Gaussian distribution at 35 weeks by plus or minus 50 g, or the standard deviation by plus or minus 25 g, the relative risk for that stratum ranges from 11 to 13. Such a fluctuation is minor in comparison with the relative risk of 21 for 1 week earlier, or the relative risk of 6.4 for 1 week later.

More generally, the result here does not depend on a Gaussian distribution as the standard. Even a rectangular distribution produces nearly the same set of relative risks. This is because the mortality curves shown in Figure 3 are roughly parallel on the log scale, making the ratio of rates between any two gestational age groups approximately constant across the spectrum of birth weight. To the degree those ratios are constant, the choice of a distribution for standardization is inconsequential.[12] The central requirement for the method of standardization used here is that all data be compared on a scale of relative birth weight (the z scale), rather than on a scale of absolute birth weight.

The present analysis highlights the importance of gestational age, and at the same time confirms a strong link between birth weight and perinatal mortality at each fixed gestational age. In his commentary on Wilcox and Russell's general approach, Peters stated that it was not possible to assess the "ultimate validity" of their model without considering gestational age.[13] Figure 3 shows that the relation between birth weight and perinatal mortality is not altered by stratifying on gestational age.

Gestational maturity has been understood for centuries to be important to infant survival. Still, gestational age tends to be slighted in contemporary US research.

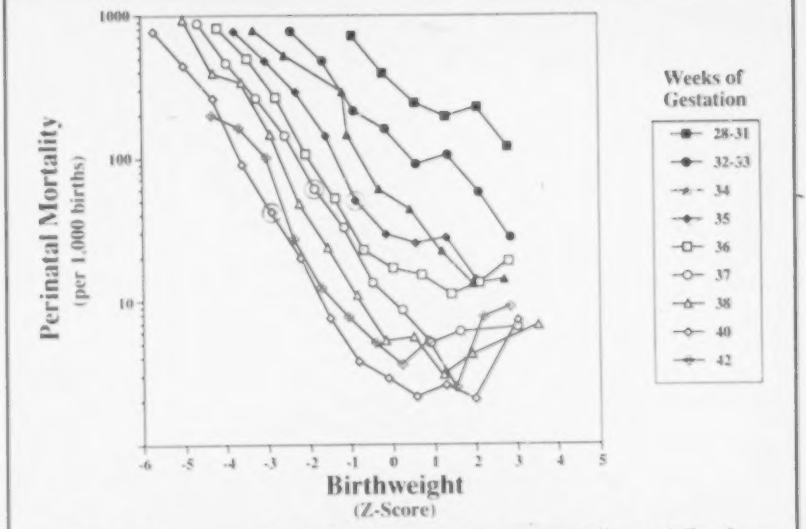


FIGURE 3—Z-score–specific mortality rates for nine representative gestational-age strata of births. (Circled points show mortality for 2250-g babies at different gestational ages. See text for discussion.)

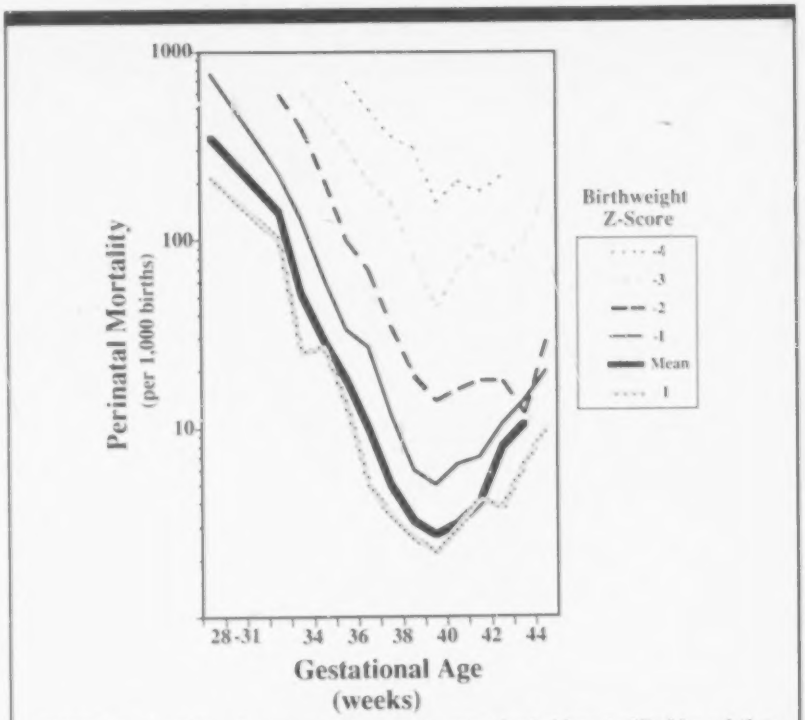


FIGURE 4—Gestational-age–specific mortality rates for babies stratified by relative size (in standard deviations from the Gaussian mean).

This neglect may be due to a lack of clarity about the role of gestational age in the causal pathway. We have shown that there are two strong and separable factors affecting perinatal survival. One is gestational age, and the other is relative birth weight at any given gestational age. A baby can benefit as much from an increase in gestational age as from an increase in its weight relative to the weights of others at the same gestational age. This benefit of additional weeks of gestational age tends to be obscured by the customary multivariate methods of "controlling" for birth weight. There is a fallacy in inferring causality from a highly predictive relationship in a statistical model. The dominance of birth weight when the ordinary analytic methods are used may have contributed to the current emphasis on low birth weight as a public health problem.[14] Interventions aimed at increasing the size of babies may have little effect on perinatal mortality. Preterm delivery appears to be as worthy a target for public health intervention as low birth weight, and may be more amenable to change. □

Acknowledgments

Clarice Weinberg and Mervyn Susser contributed insightful comments to this paper. Other helpful suggestions were provided by Pierre Buekens, Glinda Cooper, Siobán Harlow, Irva Hertz-Picciotto, Paige Hornsby, Rolv Terje Lie, Ruth Little, Walter Rogan, Andrew Rowland, David Savitz, and Beth Whelan. Figures were generated by Mike Day and Bob McConnaughey.

References

1. Wilcox AJ, Russell IT. Birthweight and perinatal mortality: III. towards a new method of analysis. *Int J Epidemiol.* 1986;15:188–196.
2. Wilcox AJ, Russell IT. Why small black infants have a lower mortality than small white infants: the case for population-specific standards for birth weight. *J Pediatr.* 1990;7–10.
3. Gruenwald P. Growth of the human fetus: I. normal growth and its variation. *Am J Obstet Gynecol.* 1966;94:1112–1119.
4. Milner RDG, Richards B. An analysis of birth weight by gestational age of infants born in England and Wales, 1967 to 1971. *J Obstet Gynecol Br Commonwealth* 1974;81:956–967.
5. Skjærven R, Wilcox AJ, Russell D. Birthweight and perinatal mortality of second births conditional on weight of the first. *Int J Epidemiol.* 1988;17:830–838.
6. Wilcox AJ, Russell IT. Birthweight and perinatal mortality: I. on the frequency distribution of birthweight. *Int J Epidemiol.* 1983;12:314–318.
7. Wilcox AJ, Russell IT. Birthweight and perinatal mortality: II. on weight-specific mortality. *Int J Epidemiol.* 1983;12:319–325.
8. Wilcox AJ. Intrauterine growth retardation: beyond birthweight criteria. *Early Hum Dev.* 1983;8:189–193.
9. Susser M, Marolla FA, Fleiss J. Birth weight, fetal age and perinatal mortality. *Am J Epidemiol.* 1972;96:197–204.
10. McCormick MC. The contribution of low birth weight to infant mortality and childhood morbidity. *New Engl J Med.* 1985; 312:82–90.
11. Skjærven R, Bakketeig LS. Classification of small-for-gestational-age births: weight-by-gestation standards of second birth conditional on the size of the first. *Pædiatr Perinat Epidemiol.* 1989;3:432–447.
12. Fleiss JL. *Statistical Methods for Rates and Proportions.* New York: John Wiley & Sons; 1973.
13. Peters TJ. Birthweight and perinatal mortality: a tale of two populations? *Pædiatr Perinat Epidemiol.* 1988;2:98–100.
14. Institute of Medicine, Committee to Study the Prevention of Low Birthweight. *Preventing Low Birthweight.* Washington, DC: National Academy Press; 1985.

Estimation of the Break-Even Point for Smoking Cessation Programs in Pregnancy

ABSTRACT

Background. Successful programs to help pregnant women quit smoking have been developed and evaluated, but formal smoking cessation programs are not a part of care at most prenatal sites. The cost of such programs may be an issue. Considering the costs of adverse maternal and infant outcomes resulting from smoking, we estimated there would be an amount of money a prenatal program could invest in smoking cessation and still "break even" economically.

Methods. A model was developed and published data, along with 1989 hospital charge data, were used to arrive at a break-even point for smoking cessation programs in pregnancy.

Results. Using overall United States data, we arrived at a break-even cost of $32 per pregnant woman. When these data were varied to fit specific US populations, the break-even costs varied from $10 to $237, with the incidence of preterm low birth weight having the most impact on the cost.

Conclusions. It may be advisable to invest greater amounts of money in a prenatal smoking cessation program for some populations. However, for every population there is an amount that can be invested while still breaking even. (*Am J Public Health.* 1992;82:383–390)

Miriam Shipp, MD, MPH, Mary S. Croughan-Minihane, PhD, Diana B. Petitti, MD, MPH, and A. Eugene Washington, MD, MSc

Introduction

Cigarette smoking has been linked with adverse pregnancy outcome since at least the 1940s[1,2] and is the most important cause of low birth weight (LBW) in developed countries.[3] However, smoking cessation during the first trimester leads to infant outcomes similar to those for nonsmokers,[4] suggesting that smoking cessation programs during early pregnancy can alter pregnancy outcomes.

Although successful programs to help pregnant women quit smoking have been developed and evaluated,[5–8] formal antismoking programs are not a part of most prenatal care. The reasons why smoking cessation programs have apparently not been implemented are not known with certainty, but the cost of such programs is probably a major issue.

A recent analysis[9] of the cost-effectiveness of smoking cessation programs in pregnancy, however, showed that over $3 was saved for every dollar spent on smoking cessation. Their cost-effectiveness was also demonstrated in a study that examined such a program in a health maintenance organization (HMO) population.[10]

These two studies looked at cost-effectiveness by assuming a fixed program cost and comparing this cost with that of the adverse medical outcomes. An alternative approach to cost-effectiveness analysis is to examine the cost of adverse medical outcomes with an intervention and compare it with the cost without such an intervention. The difference in these two costs is the amount that the program could invest in the intervention and still "break even" economically.

Our analysis takes such a perspective. We calculated the aggregate costs of medical outcomes for a hypothetical prenatal program that included a formal smoking cessation intervention and compared them with a hypothetical prenatal program that did not have such an intervention. Assuming that any program must decide how to allocate fixed resources for a group of pregnant women, who comprise both smokers and nonsmokers, we therefore arrived at an estimation of cost per pregnant woman rather than of cost per pregnant smoker in our final result. Then, after arriving at this break-even point, we examined the extent to which this amount might vary from program to program, depending on the prevalence of smoking, the success of the cessation program, the incidence of adverse pregnancy outcomes in the population in question, and the cost of medical care.

Methods

Decision Analysis Models

We constructed two decision trees—one for infant outcome (Figure 1) and one maternal outcome (Figure 2)—to compare two strategies: (1) providing a

The authors are all at the School of Medicine of the University of California at San Francisco—Miriam Shipp with the Institute for Health Policy Studies, Mary S. Croughan-Minihane and Diana B. Petitti with the Department of Family and Community Medicine, and A. Eugene Washington with the Institute for Health Policy Studies and the Department of Obstetrics, Gynecology, and Reproductive Sciences.

Requests for reprints should be sent to Dr. Mary Croughan-Minihane, Department of Family and Community Medicine, University of California, Box 0900, Rm AC-9, San Francisco, CA 94143-0900.

This paper was submitted to the journal February 22, 1991, and accepted with revisions July 10, 1991.

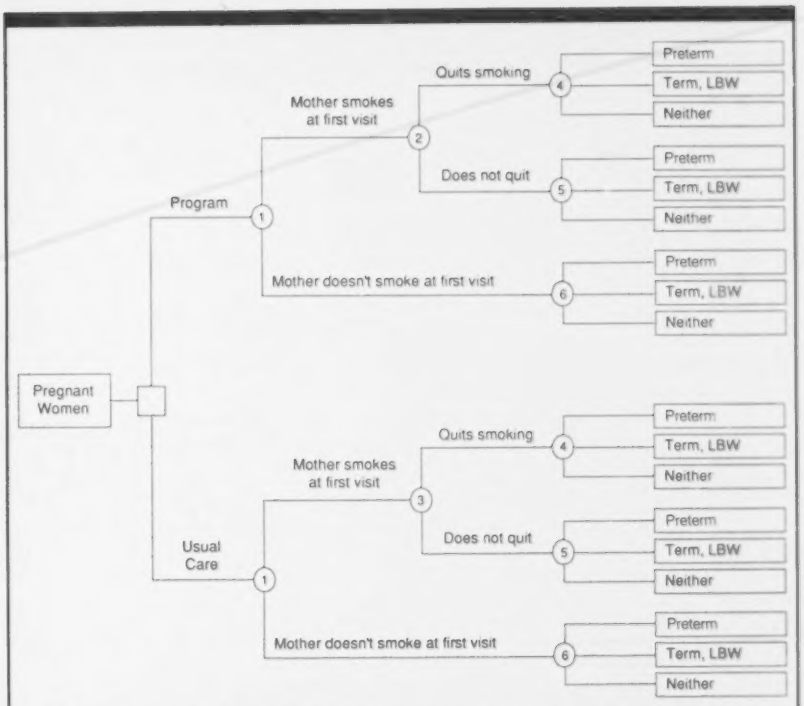


FIGURE 1—Decision tree depicting infant outcome for pregnant women provided or not provided with a formal smoking cessation program. Circles are chance nodes and square is a decision node.

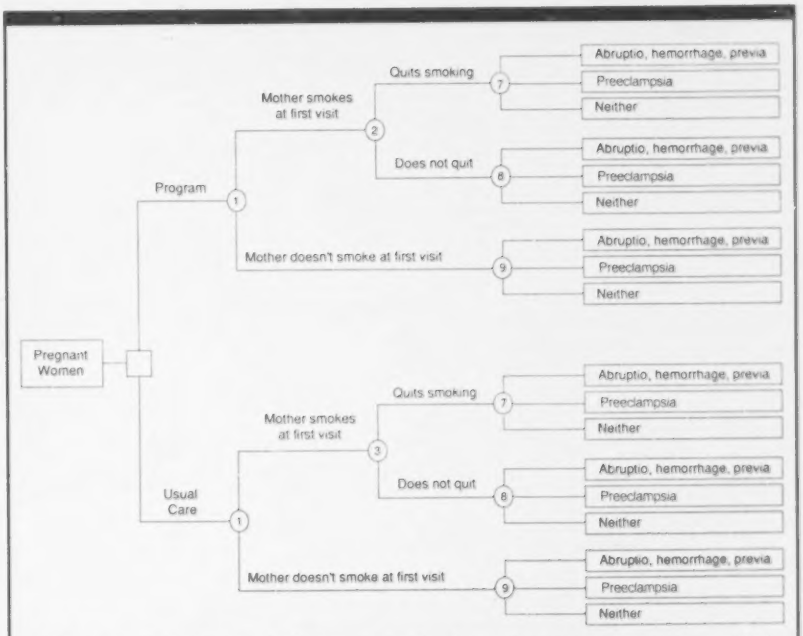


FIGURE 2—Decision tree depicting maternal outcome for pregnant women provided or not provided with a formal smoking cessation program.

formal smoking cessation program for pregnant women who smoke, or (2) not providing a formal smoking cessation program. Probabilities used in the decision model were estimated based on published data (Table 1).

Data and Assumptions

Prevalence of Smoking at First Prenatal Visit (Node 1). Population-based data on smoking prevalence were obtained from the 1985 National Health Interview Survey (published in 1988).[11] Among survey respondents, 31.8% of the women between the ages of 18 and 44 who had given birth to a child within the past 5 years smoked cigarettes during the 12 months preceding pregnancy. Because 21.2% of the women quit smoking upon learning they were pregnant, we estimated a baseline probability of 25% for women still smoking at the time of their first prenatal visit (.318–(.318) (.212) = .25). This prevalence of .25 is the figure used for all the baseline analyses. The greatest variability in smoking prevalence was for women with different education levels. Therefore, probability estimates for the sensitivity analysis were based on a comparison of women with fewer than 12 years of education (probability .39) and women with 16 or more years of education (probability .08). These estimates take into account the before-pregnancy smoking rates and the chances of women quitting upon learning of pregnancy.

Quit Rate after First Prenatal Visit in Pregnant Smokers Who Are Provided with a Formal Program (Node 2). Randomized clinical trials of smoking cessation programs allow us to estimate quit rates attributable to formal smoking cessation programs. An optimal intervention program for pregnant women was implemented and evaluated by Sexton and Hebel,[7] who achieved the highest quit rate among pregnant smokers published to date (43%). Subtracting the quit rate in the control group (20%) in that study from the quit rate in the intervention group, we estimated a quit rate attributable to the program of 23%; this figure was used for the baseline analysis. In a study of a different smoking cessation program by Windsor et al.,[6] the quit rate attributable to the program was 12% (14% intervention, 2% control), which provided the lower probability estimate for our sensitivity analysis. We obtained the upper probability estimate of 29% by taking the difference in the smoking cessation rates between the intervention (43%) and control (14%) groups for men who participated in the Multiple Risk Factor Intervention Trial (MRFIT)[12] study. Although the MRFIT study evaluated men rather than pregnant women, we used it as the high estimate of the effectiveness of a smoking cessation program because MRFIT's smoking intervention

TABLE 1—Probabilities Used in Decision Analysis

	Baseline Probability	Range of Probability	Node	Source
Prevalence of smoking at first prenatal visit	.25	.08–.39	1	NCHS, 1988[11]
Quit rate in pregnant smokers[a] who are provided with a formal program	.23	.12–.29	2	Sexton and Hebel, 1984[7] Windsor et al., 1985[6] MRFIT 1982
Spontaneous quit rate in pregnant smokers[a]	.02	.00–.06	3	Windsor et al., 1985[6] MacArthur et al., 1987[13]
Probability of LBW (<2500 g) at term (≥37 wks)				
nonsmokers	.017	.015–.031	6	NCHS 1988[11]
quit smoking after 1st visit	.017	.015–.031	4	Puffer and Serrano, 1987[14] Ounsted et al., 1985[15]
continue smoking after 1st visit	.060	.053–.109	5	McIntosh, 1984[16]
Probability of LBW (<2500 g) and preterm (<37 wks)				
nonsmokers	.035	.029–.066	6	NCHS, 1988[11]
quit smoking after 1st visit	.035	.029–.066	4	Puffer and Serrano, 1987[14] Kramer, 1987[3]
continue smoking after 1st visit	.049	.041–.092	5	McIntosh, 1984[16]
Probability of normal birth weight (≥2500 g) at term (≥37 wks)				
nonsmokers	.948	.903–.956	6	NCHS 1988[11]
quit smoking after 1st visit	.948	.903–.956	4	Puffer and Serrano, 1987[14] Ounsted et al., 1985[15]
continue smoking after 1st visit	.891	.799–.906	5	Kramer, 1987[3]
Probability of abruptio placenta, placenta previa, or antepartum hemorrhage				
nonsmokers	.009	.007–.021	9	Meyer et al., 1976[23]
quit smoking after 1st visit	.009	.007–.021	7	Reid et al., 1972[25] Cunningham et al., 1989[24]
continue smoking after 1st visit	.013	.010–.029	8	McIntosh, 1984[16]
Probability of preeclampsia of pregnancy				
nonsmokers	.058	.058–.117	9	Marcoux et al., 1989[27]
quit smoking after 1st visit	.058	.058–.117	7	Cunningham et al., 1989[24] Goplerud, 1982[26]
continue smoking after 1st visit	.024	.024–.049	8	
Probability of none of the above complications				
nonsmokers	.933	.862–.935	9	Marcoux et al., 1989[27]
quit smoking after 1st visit	.933	.862–.935	7	Cunningham et al., 1989[24] Goplerud, 1982[26]
continue smoking after 1st visit	.963	.922–.966	8	Reid et al., 1972[25] Meyer et al., 1976[23]

[a]women who smoke at the first prenatal visit.

was highly effective and may represent the maximum achievable for a formal cessation program. There is no assurance, however, that the same results could be achieved in pregnancy.

Spontaneous Quit Rate after First Prenatal Visit in Pregnant Smokers (Node 3). Spontaneous quit rates were assumed to be the same as quit rates in women in the control groups of randomized clinical trials. The baseline probability (.02) was derived from the 2% quit rate found among control group women participating in the study by Windsor et al.[6] The upper probability range (.06) was derived from a study by MacArthur et al.,[13] and the probability range of zero was chosen as an alternative, assuming that all women who are going to quit smoking do so prior to their first prenatal visit.*

*Regarding the estimates for nodes 2 and 3, we used data from randomized trials of smoking cessation programs to determine the net effect attributable to the smoking cessation program, which was estimated as the difference between the quit rates in the intervention group and the control group. This was the estimate used for the program arm of the decision tree. In the nonprogram arm of the decision tree, we also used this data on quit rates from the control groups to estimate the community spontaneous quit rate because these studies are among the few reliable sources of data on quit rates among pregnant women.

Infant Outcomes

Probability of LBW (< 2500 g) at Term (≥ 37 weeks) for Nonsmokers and Quitters (Nodes 4 and 6). Only incidence rates and relative risks were available in the literature. Thus, we used the following formula to calculate the probability of LBW at term for nonsmokers and quitters:

$$P_{NS} = \frac{I_{LBW}}{(RR_s)(Pv_s) + (Pv_{NS})},$$

where

P_{NS} = the probability of the outcome among nonsmokers and quitters,

I_{LBW} = the incidence of LBW at term for the population,

RR_S = the relative risk of LBW at term for continuing smokers,

Pv_S = the prevalence of smoking for the population, and

$Pv_{NS} = 1-Pv_S$.

The incidence rate of LBW at term (2.8%) was derived from US population-based data,[14] and the relative risk of term LBW for smokers (3.5) was obtained from a study by Ounsted et al.[15] The value used for the prevalence of smoking at the first prenatal visit was .25, and was derived as explained earlier.

Probability of LBW (< 2500 g) and Preterm (< 37 weeks) for Nonsmokers and Quitters (Nodes 4 and 6). The same formula was used to calculate the probability of LBW and preterm for nonsmokers and quitters, substituting a baseline overall incidence of preterm LBW of 3.8% and a baseline relative risk of preterm LBW of 1.4.[3,14]

Probability of LBW (< 2500 g) at Term (≥ 37 weeks) and Preterm (< 37 weeks) for Continuing Smokers (Node 5). The probability of each outcome among continuing smokers was derived from the probability of the outcome among nonsmokers and quitters and the relative risk of the outcome for continuing smokers:

$$P_S = (P_{NS})(RR_S),$$

where

P_S = the probability of the outcome among continuing smokers,

P_{NS} = the probability of the outcome among nonsmokers and quitters, and

RR_S = the relative risk of LBW for continuing smokers.

The range of relative risk estimates for continuing smokers was obtained from the Alameda County Low Birth Weight Study.[4] These were 4.5 to 5.1 for term LBW and 2.2 to 3.3 for preterm LBW.

Probability of Normal Birth Weight (> 2500 g) at Term (≥ 37 weeks) for Continuing Smokers (Nodes 4,5,6). The probability of normal birth weight at term for each smoking category was derived by subtraction:

$$P_{Norm} = 1-P_{Preterm}-P_{Term},$$

where

P_{Norm} = the probability of normal birth weight at term,

$P_{Preterm}$ = the probability of LBW and preterm, and

P_{Term} = the probability of LBW at term.

Sensitivity Analysis for Infant Outcome. Table 1 presents the range of probability estimates used in the sensitivity analysis. The first sensitivity analysis varied the prevalence of smoking at the first prenatal visit, using the lower (.08) and upper (.39) probability range values to determine the individual effect of smoking prevalence on the break-even point. Additional analyses were then conducted using the probability range estimates for each additional variable, varying only one estimate at a time. The analysis for the variable preterm LBW had values ranging from a low incidence of 3.2% to a high incidence of 7.3%, and the values for LBW at term were 2.4% and 5.0%, respectively.[14]

Because exact probabilities of infant outcomes could not be obtained from the literature and because the model appeared to be particularly sensitive to the relative risk of preterm LBW, a second sensitivity analysis examined the independent effects of varying the incidence rate for each outcome and relative risk estimate for preterm LBW. A final sensitivity analysis examined the independent effects of varying the incidence rate for each outcome and relative risk estimate for term LBW.

Maternal Outcomes

Probability of Antepartum Hemorrhage, Abruptio Placenta, and Placenta Previa for Smokers, Nonsmokers, and Quitters (Nodes 7, 8, 9). Studies[16–23] report elevated relative risks among smokers for placental abnormalities or bleeding during pregnancy. We converted these relative risks to probabilities using the formula cited earlier. An incidence of .01 for hemorrhage[24–26] and a relative risk of 1.43 for hemorrhage in smokers[23] were used in the baseline formula, and we obtained the probability of hemorrhage in continuing smokers (node 8) by the same method we used for infant outcomes.

Probability of Preeclampsia in Smokers, Nonsmokers, and Quitters (Nodes 7, 8, 9). The formula cited earlier was used to calculate the probability of preeclampsia in nonsmokers and quitters. The incidence rate of preeclampsia for pregnant women in the United States (.05) was obtained from studies cited in general obstetrics texts.[24,25] The baseline relative risk of preeclampsia for smokers (.42) was obtained from a recent case-control study,[27] which obtained a relative risk value similar to that obtained in older studies.[19] The reason for the lower relative risk of preeclampsia in smokers is uncertain, but it has been theorized that nicotine might inhibit the potent vasoconstrictor thromboxane, a substance found to be increased in studies in preeclamptic women.[28,29]

For the probability of preeclampsia in continuing smokers (node 8), we multiplied the relative risk in smokers (.42) by the probability of preeclampsia in nonsmokers and quitters (P_{NS}).

Probability of Births Not Complicated by Placenta Previa, Placenta Abruptio, Hemorrhage, or Preeclampsia (Nodes 7, 8, 9). The probability of a pregnancy not complicated by any of the above diagnoses was derived by subtraction.

Sensitivity Analysis for Maternal Outcomes. For the range of probabilities cited in Table 1, we varied the incidence rates of hemorrhage and preeclampsia through the ranges cited in the literature (.008 to .023 for hemorrhage and .05 to .10 for preeclampsia) while maintaining the relative risks for continuing smokers.

Additional sensitivity analyses varied, one at a time, the prevalence of smoking at the first prenatal visit, the probability of quitting smoking with a smoking program in place, and the probability of quitting smoking without a program in place.

Costs

Because true cost data cannot be obtained and results of this analysis are of most interest to third-party payors, charges were used as a proxy for costs. We included in our analysis only the direct medical charges for maternal care at delivery and for hospital care for newborns. With no national data available at the level of detail needed for this analysis, we approximated charges based on 1989 hospital and physician charges from two San Francisco Bay Area hospitals with perinatal databases. Because exact obstetric and newborn charge data are unavailable to compare San Francisco hospital charges with national averages, we varied the ranges of charges in our sensitivity analysis from 50% to 200% of the San Francisco charges, a range that includes most hospital charges in the United States.[30] The cost assumptions used in the model are summarized in Table 2 and were derived as described below.

Maternal Charges. We identified all patients discharged from the two study hospitals with antepartum hemorrhage or preeclampsia (ICD-9 codes 641.11, 641.21, 642.41, and 642.51)[31] between July 1 to December 30, 1989. We took a systematic random sample of women who delivered during the same period without hemorrhage or preeclampsia.

We did not include maternal hospitalization prior to delivery for any of the diagnoses, given that reliable data were not available on this parameter. However, predelivery charges are unlikely to make a major contribution to total charges. We obtained hospitalization bills for 94% of the women identified (124 of 132), and to each bill we added a standard obstetrician charge for the delivery of an infant.

For our mean baseline value for both infant and maternal charges, we eliminated the highest and lowest values, presuming them to be outliers.

Infant Charges. All infants diagnosed as term LBW born in the study period were identified using the obstetrical database of each hospital (n = 31). For infants who were preterm LBW and term not LBW, we took a systematic random sample of discharges from the same period (n = 30 for each diagnosis). For all these infants, we obtained hospitalization bills for 96% (87 of 91).

We approximated physician charges for preterm and term LBW infants by taking 20% of hospital charges for each infant, a number found in an earlier study,[32] to be the average physician charge for one of the same Bay Area hospitals used in our study. For the infants who were not LBW, we used a standard charge billed by the pediatric group of each institution for care of a normal newborn.

Sensitivity Analysis for Costs. Because hospital and physician charges for the same diagnosis may vary widely due to regional and institutional characteristics, we performed a sensitivity analysis by varying the amount charged for each diagnosis from 50% to 200%.

Results

Table 3 summarizes the estimated break-even cost per pregnant woman of a program for smoking cessation in pregnancy, based on various outcomes and population characteristics. Using our baseline assumptions and considering only infant outcomes, we estimate that $35 per pregnant woman is the largest amount the program could cost without exceeding the cost of care for LBW infants later on.

TABLE 2—Hospital and Physician Mean Charges for Selected Diagnoses Associated with Smoking in Pregnancy: San Francisco, 1989

	Number of Patients	Mean Charges ($)
Maternal		
Hemorrhage	35	11 056
Preeclampsia	51	8891
No hemorrhage or preeclamp	38	6794
Infant		
<37 wks, <2500 g	28	43 755
≥37 wks, <2500 g	29	4978
≥37 wks, ≥2500 g	30	2738

TABLE 3—Estimated Break-Even Cost[a] per Pregnant Woman of a Program for Smoking Cessation in Pregnancy, by Outcomes Considered and Risk Characteristics of Population in Terms of LBW

	Outcomes Considered ($)	
Risk of LBW	Infant Only	Infant and Maternal
Baseline	35	32
Low incidence preterm LBW[b]	28	25
High incidence preterm LBW[c]	67	64
Low incidence Term LBW[d]	30	27
High incidence Term LBW[e]	64	61

[a]number of dollars per pregnant woman that could be invested in smoking cessation program with a net expenditure of zero after taking into account costs due to changes in cost of medical care for other pregnancy outcomes.
[b]Using the lower probability estimate of preterm LBW for nonsmokers and quitters (.029).
[c]Using the upper probability estimate of preterm LBW for nonsmokers and quitters (.066).
[d]Using the lower probability estimate of term LBW for nonsmokers and quitters (.015).
[e]Using the upper probability estimate of term LBW for nonsmokers and quitters (.031).

This amount is the estimated break-even cost. Using our baseline assumptions and considering both maternal and infant outcomes, the break-even cost of a program for smoking cessation during pregnancy is estimated to be $32 per pregnant woman. This is $3 less than the cost considering only infant outcomes because the higher risk of preeclampsia in nonsmokers results in an increase in expenditures.

Considering only infant outcomes, estimates of the break-even cost of such a program range from $28 to $67, depending on the assumed background incidence of preterm LBW in a population, and from $30 to $64, depending on the assumed incidence of term LBW (Table 3). When both maternal and infant outcomes are considered, the estimates of break-even cost are not much different.

Using the baseline assumptions, estimates of the break-even cost vary, depending on the percentage of women who smoke at the first prenatal visit. These estimates range from $12 per woman in a population with 10% smokers to $45 per woman in a population with 35% smokers (Figure 3).

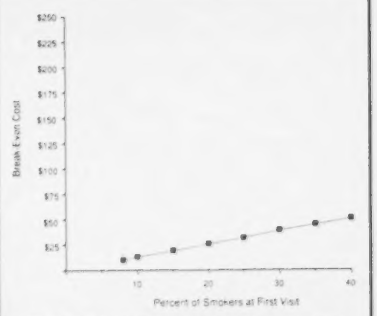


FIGURE 3—Estimated break-even cost of a program for smoking cessation during pregnancy, by prevalence of smoking at the first prenatal visit, considering both infant and maternal outcomes and using baseline assumptions for all other variables.

The estimated break-even cost strongly depends on assumptions about

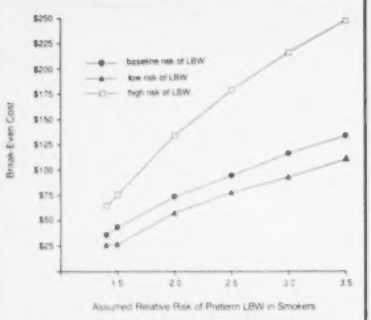


FIGURE 4—Estimated break-even cost of a program for smoking cessation during pregnancy for different assumptions about the relative risk of preterm LBW and several estimates of the background incidence of preterm LBW, considering both infant and maternal outcomes and using baseline assumptions for all other variables.

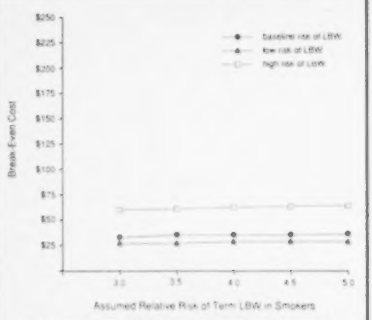


FIGURE 5—Estimated break-even cost of a program for smoking cessation during pregnancy for different assumptions about the relative risk of term LBW and several estimates of the background incidence of preterm LBW, considering both infant and maternal outcomes and using baseline assumptions for all other variables.

the relative risk of preterm LBW in continuing smokers (Figure 4). If the relative risk of preterm LBW in smokers is 3.0, as has been reported in one study of an urban Black population with a high background incidence of preterm LBW,[4] the break-even cost is $215 per woman. The estimated break-even cost does not depend on assumptions about the relative risk of term LBW in smokers (Figure 5).

TABLE 4—Estimated Break-Even Cost for Several Assumptions about the Cost of Various Outcomes, under Baseline Assumptions, Considering Both Maternal and Infant Outcomes

	Infant	Maternal	Net Cost ($)
Baseline	35	(−3)	32
Preterm LBW 50%	19	(−3)	16
Preterm LBW 200%	67	(−3)	64
Term LBW 50%	30	(−3)	27
Term LBW 200%	46	(−3)	43
Preeclampsia 50%	35	(+5)	40
Preeclampsia 200%	35	(−19)	16
Hemorrhage 50%	35	(−4)	31
Hemorrhage 200%	35	(−1)	34

As might be anticipated, assumptions about the effectiveness of the smoking cessation program influence estimates of the break-even cost. If the rate of smoking cessation were as low as 10% as a result of the program, the break-even cost drops to $12 per woman; however, if the program results in a net rate of smoking cessation as high as 29%, the break-even cost is $41.

Assumptions about the cost of care for a preterm LBW infant affect the estimate of the break-even cost more than assumptions about the cost of medical care for term LBW infants or for either of the maternal complications whose risk is increased in smokers (Table 4). No reasonable estimate of the cost of medical care for preeclampsia results in a net increase in expenditures when both infant and maternal outcomes are considered.

Discussion

Our analysis shows that, even when considering only direct medical charges related to hospitalization at delivery, a smoking cessation program that invests $32 for every pregnant woman seen in the facility would still break even economically in the short term. The number of dollars that could be invested in a smoking cessation program and still allow the program to break even is greater for populations with higher incidences of LBW, greater relative risks of LBW in smokers, and higher prevalence of smoking than the national average. However, at every level of LBW and of smoking prevalence found in the literature, money could be invested in a smoking cessation program and the program would still break even.

It should be noted that, consistent with the program perspective of our analysis described earlier, our "per patient" cost estimates apply to a cost for all patients, combining smokers and nonsmokers. To obtain the amount that could be invested per smoker, the "per patient" cost is divided by the prevalence of smoking in the population (e.g., in a baseline population with a smoking prevalence of 25%, $32/.25, or $128.00, could be spent for each smoker).

The largest cost associated with medical care of pregnant smokers and their infants is for infant care and not for care of maternal complications. There are three reasons for this. First, adverse infant outcomes are more common in smokers than are adverse maternal outcomes. Second, charges for adverse infant outcomes are much higher than those for maternal complications. Third, preeclampsia is less common in smokers than in nonsmokers.

The estimate of a break-even cost was most sensitive to assumptions about the incidence of preterm LBW and the relative risk for preterm LBW for smokers. For instance, in a population studied recently with a high incidence of and high relative risk for preterm LBW, the break-even cost was $237.[4] Although the prevalence of smoking at the first prenatal visit and the success rate of the formal smoking cessation program also affected the cost analysis, within the range of probabilities in the literature they did not produce as much variability in the break-even cost.

The break-even cost is moderately influenced by the quit rate for smokers attributable to the program, a rate that has varied considerably for programs studied.[6-8,12] Our baseline analysis is derived from a fairly optimistic quit rate of 23% obtained from an intensive intervention program for pregnant women that included one home visit, one telephone call per month, and twice-per-month mailings of smoking cessation literature.[7] More modest quit rates of 12% to 13.6% were

obtained by programs that included only short health education sessions and printed materials.[6,8] These more modest programs cost only $7.13 to $11 per pregnant smoker,[8,33] an amount below our calculated break-even cost when we substituted these quit rates into our analysis.

There are several important limitations to our study. First, we included only medical charges for the hospital admission for delivery and in-hospital charges for the neonate after delivery; therefore, our estimates are conservative. Second, we used local charge data because they are precise for LBW diagnoses. However, in our sensitivity analysis, we varied charges for each diagnosis from 50% to 200% of the mean charge, figures that include the charges for services at most US hospitals.[30] Last, we included only direct medical costs in our analysis. Cost analyses of other conditions often include indirect costs to society associated with years of potential life lost per premature death[34,35] and days of work lost due to illness.[34,36] Moreover, some studies have shown increased infant mortality in babies of smokers.[37,38] Still others suggest that children of smokers may have deficiencies in growth, intellectual development, and behavior[39–41] or an increased risk of otitis media.[42,43] Again, not considering these other adverse effects of smoking in pregnancy makes our analysis conservative.

A recent cost-effectiveness analysis was done by Marks et al.,[9] who hypothesized that $30 per pregnant smoker could be spent for a model smoking cessation program; with a 15% cessation rate, this would cost $23 505 300 nationally per year, or $4000 per LBW baby averted. The study compared this figure with NICU costs for LBW babies of smokers and surmised that the program would save $3.31 for every $1 invested. This ratio would increase to over $6 saved for every $1 spent if the costs of long-term care for infants with disabilities secondary to LBW were included. Although this study differed from the present analysis in several ways, the major difference was one of perspective: their analysis calculated the costs for a smoking cessation program nationally and compared them with the costs of care for LBW infants; our study takes a program perspective.

Importantly, our analysis includes a way of determining which variables the analysis is most sensitive to, and it provides a way for planners to input their own values for these population variables. Both our study and the Marks et al. study, as well as the HMO study mentioned earlier,[10] conclude that smoking cessation programs in pregnancy are cost-effective, whether from a national or a program perspective. Our analysis adds a dimension that allows program planners to determine how much to spend on these programs.

Program planners and administrators can use our analytic framework along with their own data to help decide how much to invest in a smoking cessation program. However, planners may want to consider factors other than the immediate break-even cost in their decision to include a smoking cessation program in their prenatal care. Considering the other long-range adverse medical, psychological, and societal outcomes of smoking, planners and society should be willing to invest in smoking cessation programs for pregnant women, regardless of their direct benefit of saving money. □

Acknowledgments

This work was supported in part by a fellowship grant from the Pew Charitable Trusts and funds provided by the cigarette and tobacco surtax fund of the State of California through the Tobacco-Related Disease Research Program of the University of California. The authors also acknowledge Marcus McCrory and the employees of San Francisco General Hospital for their invaluable assistance with the charge data.

References

1. Athayd E. Incidencia de abortos e mortinalidade nasoperarias da industria de fumo. *Brsil-Med.* 1948;62:237–239.
2. Simpson WJ, Linda L. A preliminary report on cigarette smoking and the incidence of prematurity. *Am J Obstet Gynecol.* 1957;73:808–815.
3. Kramer MS. Determinants of low birth weight: methodological assessment and metaanalysis. *Bulletin of the World Health Organization.* 1987;65:663–737.
4. Alameda County Low Birth Weight Study Group. Cigarette smoking and the risk of low birth weight: a comparison in Black and White women. *Epidemiol.* 1990;1:201–205.
5. Ershoff DH, Aaronson NK, Danaher BG, Wasserman FW. Behavioral, health, and cost outcomes of an HMO-based prenatal health education program. *Public Health Rep.* 1983;98:536–547.
6. Windsor RA, Cutter G, Morris J, et al. The effectiveness of smoking cessation methods for smokers in public health maternity clinics: a randomized trial. *Am J Public Health.* 1985;75:1389–1392.
7. Sexton M, Hebel JR. A clinical trial of change in maternal smoking and its effect on birth weight. *JAMA.* 1984;251:911–915.
8. Ershoff DH, Mullen PD, Quinn VP. A randomized trial of a serialized self-help smoking cessation program for pregnant women in an HMO. *Am J Public Health.* 1989;79:182–187.
9. Marks JS, Koplan JP, Hogue CJR, Dalmat ME. A cost-benefit/cost-effectiveness analysis of smoking cessation for pregnant women. *Am J Prev Med.* 1990;6(5):282–289.
10. Ershoff D, Quinn VP, Mullen PD, Lairson DR. Pregnancy and medical cost outcomes of a self-help prenatal smoking cessation program in a HMO. *Public Health Rep.* 1990;105(4):340–347.
11. National Center for Health Statistics. *Vital and Health Statistics.* Series 10, No. 163. 1988. DHHS publication PHS 88–1591.
12. Multiple Risk Factor Intervention Trial Research Group. Multiple risk factor intervention trial: risk factor changes and mortality results. *JAMA.* 1982;248:1465–1477.
13. MacArthur C, Newton JR, Knox EG. Effect of anti-smoking health education on infant size at birth: a randomized controlled trial. *Br J Obstet Gynaecol.* 1987;94:295–300.
14. Puffer RR, Serrano CV. *Patterns of Birth Weight.* Washington, DC: Pan American Health Organization, World Health Organization; 1987. Scientific Publication 504.
15. Ounsted M, Moar VA, Scott A. Risk factors associated with small-for-dates and large-for-dates infants. *Br J Obstet Gynaecol.* 1985;92:226–232.
16. McIntosh ID. Smoking and pregnancy: attributable risks and public health implications. *Can J Public Health.* 1984;75:141–148.
17. Lowe CR. Effect of mothers' smoking habits on birth weight of their children. *Br Med J.* October 1959;10:673–676.
18. Underwood P, Hester LL, Luffitte T, Greg KV. The relationship of smoking to the outcome of pregnancy. *Am J Obstet Gynecol.* 1965;91:270–276.
19. Underwood PB, Kesler KF, O'Lane JM, Callagan DA. Parental smoking empirically related to pregnancy outcome. *Am J Obstet Gynecol.* 1967;29:1–8.
20. Butler WR, Albermon ED. *Perinatal Problems: The Second Report of the 1958 British Perinatal Mortality Survey.* Edinburgh and London: Livingstone; 1969.
21. Andrews J, McGarry JM. A community study of smoking in pregnancy. *J Obstet Gynaecol Br Commonw.* 1972;79:1057–1073.
22. Lubs ME. Racial differences in maternal smoking effects on the newborn infants. *Am J Obstet Gynecol.* 1973;115:66–76.
23. Meyer MB, Jonas BS, Tonascia JA. Perinatal events associated with maternal smoking during pregnancy. *Am J Epidemiol.* 1976;103:464–476.
24. Cunningham GF, MacDonald PC, Gant NF. *Williams Obstetrics.* 18th ed. Norwalk, Conn: Appleton and Lange; 1989.
25. Reid DE, Ryan KJ, Benirschke K. *Principles and Management of Human Reproduction.* Philadelphia, Pa: WB Saunders Co; 1972.
26. Goplerud CP. Bleeding in late pregnancy. In: Danforth WJ, ed. *Obstetrics and Gynecology.* 4th ed. Philadelphia, Pa: Harper and Row; 1982:443–456.
27. Marcoux S, Brison J, Fabia J. The effect of cigarette smoking on the risk of pre-eclampsia and gestational hypertension. *Am J Epidemiol.* 1989;130:950–957.
28. Ylikorkala O, Makila UM. Prostacyclin

and thromboxane in gynecology and obstetrics. *Am J Obstet Gynecol.* 1985;52:318–329.
29. Ylikorkala O, Vilnikka L, Lehtovirta P. Effect of nicotine on fetal prostacyclin and thromboxane in humans. *Obstet Gynecol.* 1985;66:102–105.
30. American Hospital Association. *Hospital Statistics*. Chicago: American Hospital Association, 1990.
31. International Classification of Diseases-9-Clinical Modification, Volume 1, Annotated Edition Fourth Printing. Ann Arbor, Mich: Edwards Brothers; 1988:537–541.
32. Phibbs LS, Williams RL, Phibbs RH. Newborn risk factors and costs of neonatal intensive care. *Pediatrics.* 1981;68:313–321.
33. Windsor RA, Warner KE, Cutter GR. A cost-effectiveness analysis of self-help smoking cessation methods for pregnant women. *Pub Health Rep.* 1988;103:83–88.
34. Rice DP, Hodgson TA, Sinsheimer P, Browner W, Kopstein AN. The economic costs of the health effects of smoking, 1984. *Milbank Q.* 1986;64:489–547.
35. Morbidity and Mortality Weekly Report. Smoking-attributable mortality, morbidity, and economic costs: California, 1985. *JAMA.* 1989;261:2942–2945.
36. Swank RT, Becker DM, Jackson CA. The costs of employee smoking: computer simulation of hospital nurses. *Arch Intern Med.* 1988;148:445–448.
37. Malloy MH, Kleinman JC, Land GH, Schramm WF. The association of maternal smoking with age and cause of infant death. *Am J Epidemiol.* 1988;128:46–55.
38. Kleinman JC, Pierre MB, Madons JH, Land GH, Schramm WF. The effects of maternal smoking on fetal and infant mortality. *Am J Epidemiol.* 1988;127:274–282.
39. Butler R, Goldstein H. Smoking in pregnancy and subsequent child development. *Br Med J.* 1973;4:573–575.
40. Dunn HG, McBurney AK, Ingram S, Hunter CM. Maternal cigarette smoking during pregnancy and the child's subsequent development: I. physical growth to the age of 6½ years. *Can J Public Health.* 1976;67:499–505.
41. US Dept of Health and Human Services. *The Health Consequences of Smoking for Women: A Report of the Surgeon General.* 1983:191–244. DHHS publication 410–889/1284.
42. Reed BD, Lutz LJ. Household smoking exposure: association with middle ear infections. *Fam Med.* 1988;20:426–430.
43. Hinton AE, Buckley G. Parental smoking and middle ear effusions in children. *J Laryngol Otol.* 1988;102:992–996.

Finding Fault before Drilling: A Solution for Avoiding Radon

Dr. Irina Cech calls it a $2 solution. It's a solution, however, that could add up to a great deal of financial and health savings, not to mention saving headaches, particularly for those concerned with developing public water supplies.

Cech, professor of environmental health sciences and hydrology at the School of Public Health of The University of Texas Health Science Center at Houston, has some simple advice: to avoid a radon-contaminated well, don't drill within 4 miles from a salt dome, fault, or areas known or suspected to have uranium deposits.

Radon is a radioactive byproduct of uranium decay. Because it is odorless and colorless, there is nothing to alert well drillers or household residents unless testing occurs, Cech said.

"It is very costly to drill a well only to find out that the radon concentration is too high. By examining the area for fault lines before drilling, the problem can be avoided," she said.

Cech said uranium deposits appear primarily along the coastline of Texas, but dot other areas of the state as well. Radon deposits also collect along the flanks of salt domes, also common to the coastal areas. She said faults in the domes allow radon to "leak" upward. If an aquifer, or underground layer of earth saturated with water, happens to lie above or near a salt dome, radon may enter the water supply through the faults.

To date, Cech and her students have gathered almost 1000 water samples in Texas from public and individually owned domestic water wells and surface sources, along with wells used for livestock watering, industrial purposes, and oil production. Although radon can escape through the soil and seep through foundation cracks, in Texas it is primarily found in the water supply, Cech explained.

"Concentrations of radon were particularly high in public water wells along the Gulf Coast and in parts of west-central Texas," she said, explaining that the same area has one of the highest rates for respiratory cancer mortality in the nation.

"What we found was very interesting. When wells are developed around faults, a greater likelihood exists for encountering radon," she said.

The maximum acceptable limit of radon concentration in public drinking water that the Environmental Protection Agency has proposed is 300 pCi/L. According to the EPA, the lifetime risk of developing lung cancer from household water that contains 10 000 pCi/L of radon is roughly 3 to 13 in 1000.

"About 50% of the public water supplies measured by our group exceed the recommended standards. The highest concentration, 22 000 pCi/L, was found in southwest Texas near Refugio. In Houston, the highest observed was 6000 pCi/L," she said.

"This is of concern but should not be cause for alarm since radon levels are correctable," she said, explaining that public utilities will have until 1996 to meet the EPA requirements. Aerating the water supply would allow radon to escape before the water is piped to homes and businesses.

Radon in water is a problem only when the chemical is released from the water and enters household air. As one breathes, the radon breaks down further, causing small bursts of energy that can damage lung tissue. The chemical is released, for example, when people wash dishes and clothes, take showers, and flush toilets.

"Radon is more likely to be present in homes that are heavily insulated and airtight, which is extremely common in Texas. Drinking the water is not considered as much of a hazard as breathing the radon," she said. Economical ways to reduce risk from radon are to increase air flow in and out of the house and to refrain from smoking (scientific evidence indicates that smoking may increase the risk of exposure to radon).

Because radon appears primarily along the coastal borders, it is believed that similar situations exist in other parts of the world.

Low Birth Weight in Relation to Multiple Induced Abortions

ABSTRACT

Background. Most studies report that a single induced abortion does not increase risk for delivering a low birth weight infant in a subsequent pregnancy. However, the effect of multiple abortions has not been adequately evaluated.

Methods. This relationship was studied in 6541 White women who delivered their first child between 1984 and 1987. We compared the frequencies of low birth weight (less than 2500 g) among infants born to 1999 women without prior induced abortion and 1999 women with one abortion with the frequencies of low birth weight among infants born to women with two (n = 1850), three (n = 520), and four or more (n = 173) prior induced abortions.

Results. After adjustment for confounding variables, we found no linear relationship in risk of low birth weight among women with one (relative risk [RR] = 1.2, 95% confidence interval [CI] = 0.9–1.5), two (RR = 1.5, 95% CI = 1.1–2.0), three (RR = 1.3, 95% CI = 0.8–1.9), or four or more (RR = 1.6, 95% CI = 0.9–2.9) prior induced abortions.

Conclusions. These findings confirm earlier reports of little or no evidence of harmful effects on birth weight by one or by two or more induced abortions. We further report that risk is not significantly elevated even in women with three, four, or more prior terminations of pregnancy when compared with women with one or two abortions. (*Am J Public Health.* 1992;82:391–394)

Margaret T. Mandelson, MPH, Christopher B. Maden, PhD, and Janet R. Daling, PhD

Introduction

Since 1986, approximately 1.6 million induced abortions have been performed in the United States each year.[1] As the vast majority of women who undergo abortion eventually wish to deliver, it is important to identify possible deleterious effects of induced abortion on future pregnancy outcome. Many studies report that a single induced abortion does not increase a woman's risk for delivering a low-birth-weight infant in a subsequent pregnancy.[2–6] However, the effect of multiple induced abortions is less clear.[7] Multiple induced abortions may increase the likelihood of pregnancy loss or of preterm delivery resulting in a low-birth-weight infant in two general ways: cervical trauma from forced mechanical or rapid dilatation during the abortion procedure, or cervical and uterine adhesions due to curettage or infection.

Most studies of the risks of induced abortion on future reproductive outcome were conducted before a large proportion of women had the opportunity to have two, three, or more induced abortions. We report results from a population-based study of risk of low birth weight in women with multiple induced abortions prior to their first birth.

Methods

The Washington State birth certificate uses a check-box format to record information on demographic characteristics, complications of labor and delivery, and birth outcome. Information on history of prior induced abortions—specifically, the number of prior induced abortions and the date of the last induced or spontaneous abortion—was added to the birth certificate in 1984.

Women without a prior live birth who resided in one of the three most urban counties of Washington State (King, Pierce, and Spokane) and who delivered a live-born singleton infant between January 1, 1984, and December 31, 1987, were eligible for study. Because only a small proportion of Washington residents are non-White, this analysis was restricted to White women. Two thousand women without a history of induced abortion and 2000 women with one prior abortion were randomly sampled from the population using the birth certificate tapes. All eligible women with two (n = 1852), three (n = 522), and four or more (n = 173) induced abortions prior to their birth were identified. Six women were excluded for delivering infants with birth weights considered out of range (less than 450 g [n = 4], greater than 7000 g [n = 2]). The total study population consisted of 6541 White women.

Unconditional logistic regression[8] was used to provide parameter estimates of the risk of low birth weight (less than 2500 g) associated with one, two, three, and four or more prior induced abortions

At the time of the study, all authors were with the Division of Public Health Sciences at the Fred Hutchinson Cancer Research Center and with the Department of Epidemiology at the University of Washington in Seattle. Margaret Mandelson is with the Group Health Cooperative of Puget Sound and the Department of Epidemiology at the University of Washington; Christopher Maden is with the Seattle-King County Department of Public Health.

Address reprint requests to Margaret T. Mandelson, MPH, Group Health Cooperative of Puget Sound, Center for Health Studies, Department of Preventive Care, 1730 Minor Ave, Suite 1600, Seattle, WA 98101-1448.

This paper was submitted to the Journal November 19, 1990, and accepted with revisions July 17, 1991.

TABLE 1—Selected Characteristics of Women with None, One, Two, Three, and Four or More Prior Abortions

Variable	Number of Prior Abortions									
	None		One		Two		Three		Four or More	
	No.	%	No.	%	No.	%	No.	%	No.	%
Total	1999	100.0	1999	100.0	1850	100.0	520	100.0	173	100.0
Maternal age (years)										
≤19	322	16.1	223	11.2	113	6.1	14	2.7	3	1.7
20–24	685	34.3	688	34.4	572	30.9	145	27.9	44	25.4
25–29	600	30.0	620	31.0	658	35.6	190	36.5	65	37.6
30+	388	19.4	467	23.4	506	27.4	171	32.9	61	35.3
Unknown	4	0.2	1	0.1	1	0.1	0	0.0	0	0.0
Marital status										
Married	1589	79.5	1519	76.0	1339	72.4	359	69.0	114	65.9
Unmarried	409	20.5	479	24.0	510	27.5	161	31.0	58	33.5
Unknown	1	0.1	1	0.1	1	0.1	0	0.0	1	0.6
Smoking during pregnancy										
Yes	360	18.0	562	28.1	574	31.0	153	29.4	72	41.6
No	1588	79.4	1376	68.8	1212	65.5	346	66.5	92	53.2
Unknown	51	2.6	61	3.1	64	3.5	21	4.0	9	5.2
Income ($)										
<15 000	115	5.8	89	4.5	54	2.9	14	2.7	7	4.0
15–29 999	1376	68.8	1356	67.8	1273	68.8	341	65.6	116	67.0
30 000+	258	12.9	304	15.2	298	16.1	91	17.5	34	19.7
Unknown	250	12.5	250	12.5	225	12.2	74	14.2	16	9.3
Trimester prenatal care began										
First	1602	80.1	1600	80.0	1423	76.9	387	74.1	127	73.4
Second	262	13.1	286	14.3	298	16.1	103	19.7	34	19.7
Third or none	81	4.1	60	3.0	69	3.7	18	3.5	7	4.0
Unknown	54	2.7	53	2.7	60	3.2	14	2.7	5	2.9

Note.—Percentages may not add up to or may exceed 100% due to rounding.

relative to no abortion history. Variables entered into the model as confounders included maternal age (19 years and younger, 20–24, 25–29, 30 years and older), marital status (married, not married), smoking during pregnancy (yes, no, unknown), median family income based on census tract of residence (under $15 000, $15 000–29 999, $30 000–44 999, $45 000 or more, unknown), and trimester when prenatal care began (first, second, third, or no prenatal care). Tests for linear trend in the logistic model were obtained by categorizing the exposure and entering the scored variable as continuous.[8] To separate the effect of multiple induced abortions on the risk of low birth weight from the risk of ever having had an abortion, the test of trend used one abortion as the baseline category.

Results

Women with multiple induced abortions tended to be older than women with one or no prior abortion (Table 1). Marital status was also associated with frequency of prior termination of pregnancy: a higher proportion of women with four or more induced abortions was unmarried (33.5%) compared with women with one abortion (24.0%) or no history of abortion (20.5%). Women with four or more induced abortions were more likely to smoke during pregnancy (41.6%) compared with women with one abortion (28.1%) or no prior abortions (18.0%).

Consistent trends between history of induced abortion and median census tract family income or trimester when prenatal care began were not observed (Table 1). Women with four or more prior abortions were less likely to receive prenatal care during the first trimester of pregnancy than women with no prior induced abortions (73.4% and 80.1%, respectively); however, the frequency of prenatal care beginning in the third trimester or the circumstance of no prenatal care at all did not vary appreciably according to abortion history.

Risk factors for low birth weight were strongly associated with history of multiple induced abortions. For example, 3.8% of women 30 years of age and older had four or more induced abortions prior to their first birth, compared with 3.0% among women 25–29 years of age, 2.1% among women 20–24 years of age, and 0.4% among women 19 years of age and younger (Table 1). Similarly, women who smoked during pregnancy were more than twice as likely to have had multiple abortions (four or more) compared with nonsmoking women (4.2% and 2.0%, respectively).

The crude proportion of infants with a birth weight of less than 2500 g varied between 4.4% among those born to women without a prior induced abortion and 9.6% among those born to women with four or more prior abortions (Table 2). After adjustment for maternal age, marital status, smoking during pregnancy, socioeconomic status as measured by median census tract family income, and trimester when prenatal care began, there was an indication that women with a history of one or multiple induced abortions prior to first delivery were at somewhat elevated risk (20% to 60%) for having a low-birth-weight infant compared with women with no such history. Smoking during pregnancy was an important con-

founder in these data, and adjustment for this variable reduced estimates of relative risk (RR) by approximately 15%. Most of the estimates were consistent with chance, and none was significantly different from another. Additional adjustment for method of delivery of the current pregnancy and for number of prenatal visits did not alter these findings. A trend of increasing risk of low birth weight with an increased number of prior abortions was not detected ($\chi^2_1 = 2.32, p = .13$).

TABLE 2—Standardized Relative Risks of Low Birth Weight (<2500 g) According to Abortion History

Number of Induced Abortions	N[a,b]	% Low Birth Weight	RR	95% CI
None	1941	4.4	1.0[c]	—
One	1944	5.7	1.2	0.9–1.5
Two	1788	7.7	1.5	1.1–2.0
Three	506	7.1	1.3	0.8–1.9
Four or more	167	9.6	1.6	0.9–2.9

Trend test (relative to one abortion): $\chi^2_1 = 2.32$, $P = .13$.
Note.—RR = relative risk, CI = confidence interval.
[a]Excludes women with missing data in key variables: age, marital status, and trimester prenatal care began.
[b]Adjusted for age (≤19, 20–24, 25–29, 30+), smoking during pregnancy (yes, no, unknown), marital status (married, unmarried), family income (<$15 000, 15 000–29 999, $30 000–44 999, $45 000 or greater, unknown), and trimester when prenatal care began (first, second, third, or no prenatal care).
[c]Reference category.

Discussion

Several methodological issues associated with our study may have affected our findings. This analysis used information routinely collected on the Washington State birth certificate and is limited by the accuracy of that data. The check-box format on the birth certificate has improved the accuracy of reporting pregnancy complications,[9] but the reliability of other information is less certain. It is possible that underreporting of induced abortion may have occurred. Misclassification of this exposure, particularly among women categorized as having no prior abortion history, would bias our results toward little or no association. Previous studies have questioned whether women will admit to having had an induced abortion.[10–12] However, these studies were conducted during the previous decade, when women might have been more reluctant to discuss induced abortion and when some of the abortions may have been illegally obtained. Furthermore, a recent study conducted in the same geographic area[13] found that self-reported history of induced abortion obtained through personal interview agreed closely with medical chart review.

Analyzing computerized birth certificate records may be one of the few ways to study this issue using a population-based design, but such records provide limited information on confounding factors. For example, information such as maternal education, alcohol use, and illicit drug use during pregnancy is not routinely collected. Additionally, variables on the certificate, such as average family income based on census tract of residence and dichotomized smoking information, may provide inadequate statistical control for the effect of socioeconomic status and smoking behavior. Residual confounding may be present, and our estimates of relative risk may be spuriously high.

There is evidence to suggest that the frequency of late complications may vary by the abortion procedure used.[14–15] Abortion method and gestational age at termination were not known for the women in our study. It is likely, however, that most of the abortions were performed early in pregnancy and by vacuum aspiration, given that more than 95% of induced abortions performed in the United States in recent years use this method and approximately 89% are conducted during the first trimester.[16]

Among the 94 939 White women who delivered their first live-born infant in Washington State between 1984 and 1987, 16.2% had terminated one or more prior pregnancies with an induced abortion. Although most of these women had only had one prior abortion, 3.4%, 1%, and 0.3% had, respectively, two, three, and four or more pregnancies terminated. The rate of low birth weight among infants born to women with one or more prior induced abortions was higher than it was among infants born to women with no prior induced abortions. However, after adjusting for the confounding effects of maternal age, marital status, smoking, family income, and prenatal care, we found very little, if any, increased risk associated with multiple induced abortions. Most of the estimates were consistent with chance and were of a magnitude that might be explained by residual confounding. We were unable to detect a significant linear trend in risk with an increasing number of prior induced abortions.

In light of the overall lack of association reported here between previous induced abortions and low birth weight, a consideration of study power is warranted. This study could have detected a 1.4-fold elevation in risk of low birth weight among women with one prior abortion compared with women with no prior abortions at $\alpha = .05$ and $1-\beta = .80$. Comparing women with two, three, and four or more prior abortions, the same power could have detected relative risks of 1.4, 1.7, and 2.2, respectively.

The association between induced abortion and poor reproductive outcome has been well studied and reviewed;[7] however, few studies have evaluated the specific effects of multiple induced abortions. Further, no previous study has had adequate numbers of women with two, three, or more induced abortions prior to a first birth with which to estimate the risk of low birthweight associated with each of these higher frequencies of induced abortion.

Our findings closely agreed with and extended reports by other investigators, whose data have been limited to the more general grouping of two or more prior abortions. Most recently, Bracken and coworkers[4] report no evidence for a causal relationship between one (RR = 0.86) or two or more (RR = 1.14) prior abortions and low birth weight in women delivering their second or third pregnancy. Linn et al.[5] report that women with one prior abortion have a risk of 0.74 to 1.17 for having a low-birth-weight infant relative to that of women who had not undergone an induced abortion, and that women with two or more prior abortions have a risk of 0.88 to 1.79. Our results also confirm earlier findings reported by Daling and Emanuel[6] who observed no elevation in risk of low birth weight associated with one or more induced abortions.

However, these results do not agree with a study by Seidman et al.,[17] which found an excess risk of low birth weight associated with one, two, and three or more prior abortions. Based on a series of

deliveries occurring in Israel between 1974 and 1976, women with one or more prior abortions were found to have 1.3 times the rate of low birth weight relative to women with no previous abortions. Among women delivering their first child, the unadjusted percentages of low birth weight for women with one, two, and three or more prior abortions are 1.4 to 2.2 times higher than the frequencies found in our data. These conflicting findings may be due to differences in abortion procedure and in method of cervical dilatation. Induced abortions conducted in Israel between 1974 and 1976 were generally performed by dilatation and curettage or, for late abortion, intrauterine saline injections. In contrast, the majority of abortions in our sample, as previously noted, were most likely performed by vacuum aspiration during the first trimester.

An association between multiple induced abortions and very low birth weight has been reported.[18] Australian perinatal records of births between 1982 and 1983 show that women with two or more induced abortions are 3.6 times more likely to deliver an infant weighing less than 1000 g relative to women with no abortion history. Information on confounding factors such as smoking and socioeconomic status was not recorded on these perinatal records, and analysis of risk is based on crude percentages of very low birth weight. Few women in our sample (n = 36) delivered infants of very low birth weight (less than 1500 g), and it was not possible to estimate risk of this outcome independently of the effects of important confounding factors.

These results apply only to White women delivering their first child and to the risk of only one adverse pregnancy outcome. They confirm earlier findings that show little or no evidence of harmful effects on birth weight by one or by two or more induced abortions, and we further report that this risk is not significantly elevated even in women with three, four, or more prior terminations of pregnancy when compared with women with one or two abortions. Over half of all abortions conducted in the United States during the study period were performed on women without a prior live birth.[16] The results of the present study together with those of other reports suggest that, if there is any increase in risk for having a low-birth-weight infant in the first delivery following multiple induced abortions, it is likely to be modest. □

Acknowledgments

An earlier version of this paper was presented at the 115th meeting of the American Public Health Association in New Orleans, LA, October 18–22, 1987.

References

1. Henshaw SK, Van Vort J. Abortion services in the United States, 1987 and 1988. *Fam Plann Perspect.* 1990;22:102–108.
2. Frank PI, Kay CR, Lewis TLT, Parish S. Outcome of pregnancy following induced abortion: report from the joint study of the Royal College of General Practitioners and the Royal College of Obstetricians and Gynaecologists. *Br J Obstet Gynaecol.* 1985;92:308–316.
3. Pickering RM, Forbes JF. Risks of preterm delivery and small-for-gestational-age infants following abortion: a population study. *Br J Obstet Gynaecol.* 1985;92:1106–1112.
4. Bracken MB, Hellenbrand KG, Holford TR, Bryce-Buchanan C. Low birthweight in pregnancies following induced abortion: no evidence for an association. *Am J Epidemiol.* 1986;123:604–613.
5. Linn S, Schoenbaum SC, Monson RR, Rosner B, Stubblefield PG, Ryan KR. The relationship between induced abortion and outcome of subsequent pregnancies. *Am J Obstet Gynecol.* 1983;146:136–140.
6. Daling JR, Emanuel I. Induced abortion and subsequent outcome of pregnancy in a series of American women. *N Engl J Med.* 1977;297:1241–1245.
7. Hogue CJR, Cates W Jr, Tietze C. The effects of induced abortion on subsequent reproduction. *Epidemiol Rev.* 1982;4:66–94.
8. Breslow NE, Day NE. *Statistical Methods in Cancer Research: Vol I. The Analysis of Case-Control Studies.* Lyon, France: International Agency for Research on Cancer; 1980. IARC Scientific Publications 32.
9. Frost F, Starzyk P, George S, McLaughlin JF. Birth complication reporting: the effect of birth certificate design. *Am J Public Health.* 1984;74:505–506.
10. Park TK, Straus LT, Hogue CJR, Kim IS. Previous experience of induced abortion as a risk factor for fetal death and preterm delivery. *Int J Gynaecol Obstet.* 1984;22:195–202.
11. Richardson JA, Dixon G. Effects of legal termination on subsequent pregnancy. *Br Med J.* 1976;1:1303–1304.
12. Madore C, Hawes WE, Many F, Hexter AC. A study on the effects of induced abortion on subsequent pregnancy outcome. *Am J Obstet Gynecol.* 1981;139:516–521.
13. Holt VL, Daling JR, Voigt LF, et al. Induced abortion and the risk of subsequent ectopic pregnancy. *Am J Public Health.* 1989;79:1234–1238.
14. WHO Task Force on Sequelae of Abortion. Gestation, birth-weight, and spontaneous abortion in pregnancy after induced abortion. *Lancet.* 1979;1:142–145.
15. Slater PE, Davies AM, Harlap S. The effect of abortion method on the outcome of subsequent pregnancy. *J Reprod Med.* 1981;26:123–128.
16. Koonin LM, Atrash HK, Smith JC, Ramick M. Abortion surveillance, 1986–1987. In: CDC Surveillance Summaries, June 1990. *MMWR.* 1990;39(No. SS-2):23–27.
17. Seidman DS, Ever-Hadani P, Slater PE, Harlap S, Stevenson DK, Gale R. Child bearing after induced abortion: reassessment of risk. *J Epidemiol Community Health.* 1988;42:294–298.
18. Lumley J. Very low birthweight (<1,500 g) and previous induced abortion: Victoria 1982–1983. *Aust N Z J Obstet Gynecol.* 1986;26:268–272.

Mortality, Disability, and Falls in Older Persons: The Role of Underlying Disease and Disability

ABSTRACT

Background. Falls are prevalent in older persons and can have serious consequences.

Methods. Data from the Longitudinal Study on Aging were analyzed to study the relationship between falls and both mortality and functional status in 4270 respondents age 70 and over. The effects of demographic traits, chronic conditions, and disability present at baseline were controlled for by means of multivariable analyses.

Results. Risk of death within 2 years was greater for both single fallers (crude odds ratio [OR], 1.5; 95% confidence interval [CI], 1.1–2.0) and multiple fallers (crude OR, 2.2; 95% CI, 1.7–2.8). This excess risk was dissipated when selected covariates were added to the model. No crude or adjusted association was evident between single falls and functional impairment; however, multiple falls were an independent risk factor (adjusted OR, 1.6; 95% CI, 1.2–2.0).

Conclusions. Multiple falls in older persons increase risk of functional impairment and may indicate underlying conditions that increase risk of death. (*Am J Public Health.* 1992;82:395–400)

Julie E. Dunn, MS, Mark A. Rudberg, MD, MPH, Sylvia E. Furner, MPH, PhD, and Christine K. Cassel, MD

Introduction

Falls are prevalent in older persons, and the incidence of falls increases with age. Twenty-eight percent to 33% of those aged 65 and older report having fallen over a 1-year period, and the rates approach 50% in those over 80.[1–3]

Falls can have serious consequences in older persons. People over the age of 65 have the highest mortality rate from injuries, and the largest single cause of injury mortality in this group is falls.[4] A study using the Major Trauma Outcome Study database found an 11.7% case fatality rate for falls in geriatric patients,[5] and recent data from the Dade County, Fla, community-based Study to Assess Falls Among the Elderly[6] found 2.2 deaths per 100 fall injury events that came to medical attention, not including deaths that occurred after discharge from the hospital. There is some indication that longer-term outcomes of falls may be even more serious than the short-term outcome statistics would indicate. A British study of 125 people aged 65 and older who fell in the home reported a 1-year mortality rate of 26%, compared with 6% in a control group, although none of the deaths in those who fell were reported as being from fall-related causes.[7] Nonfatal falls can also have severe consequences in the aged: the Dade County study[6] found that about half of the fall injury events occurring at home and requiring hospitalization resulted in discharge to a nursing home.

These findings raise several questions: Are falls in older persons related to increased mortality and disability? If so, to what extent are these adverse outcomes due directly to falls, rather than to factors associated with both falling and mortality or disability outcomes?

Several studies[1–3,8] show the relationship between falling and mobility disorders, diseases, sensory losses, medication use, and limitations in activities of daily living (ADLs). It has been argued that the elevated mortality associated with falling is caused by conditions that predispose toward falling, rather than from resulting trauma.[7] Also, it has been suggested that a person who may already have compromised gait and balance systems may fall when a new illness or condition, even one that seems minor or unrelated to falling per se, occurs.[9] Tinetti[10] showed that a fall risk score, derived from the number of chronic diseases present, predicted risk of falling in older persons better than did a mobility score alone. She concluded that falling "appears to result from the accumulated effect of multiple specific debilities."

Researchers have found that those who experience repeated falls differ in some respects from those who have single or occasional falls. Risk factors for single falls appear to be less robust than those for multiple falls.[8] One suggestion is that occasional falls result primarily from extrinsic factors, such as environmental haz-

Julie E. Dunn and Sylvia E. Furner are with the University of Illinois School of Public Health, Epidemiology and Biostatistics Program, Chicago, Ill. Mark A. Rudberg and Christine K. Cassel are with the Section of General Internal Medicine, Department of Medicine, University of Chicago. Christine K. Cassel is also with The Harris Graduate School of Public Policy Studies at the University of Chicago.

Requests for reprints should be sent to Julie E. Dunn, MS, University of Illinois at Chicago School of Public Health, Department of Epidemiology and Biostatistics, 2121 W Taylor St, Chicago, IL 60680.

This paper was submitted to the journal December 4, 1990, and accepted with revisions June 24, 1991.

TABLE 1—Covariates Used in Logistic Regression Analyses

Demographic	Chronic Condition	Disability
Age (1984)	Osteoporosis/hip fracture	Difficulty in
Sex	Arthritis	Bathing
Marital status	Hypertension	Dressing
Education	Heart disease	Eating
	Stroke	Transferring
	Vascular disease	Walking
	Diabetes	Going out
	Cancer	Toileting
	Visual deficit	
	Hearing deficit	
	Thinness	

ards, while repeated falls are the result of intrinsic factors, such as demographic factors, chronic disease, and disability.[9]

We have used data from the Longitudinal Study on Aging (LSOA) to examine the relationship between falls, chronic disease, disability, and mortality; the LSOA population is a larger and more diverse sample population than has been examined in previous studies of falling. Specifically, we determined the following:

1. Crude 2-year mortality rates in those who had and those who had not fallen in the year prior to the baseline (1984) LSOA survey
2. The extent to which history of falling in the past year predicts subsequent mortality, when demographic factors, chronic conditions, and functional disabilities present at baseline are controlled for
3. The extent to which history of falling predicts decreased functional ability after 2 years, when demographic variables, chronic conditions, and disabilities present at baseline are controlled for.

Methods

The LSOA Population

For this analysis we used data from the 1986 longitudinal follow-up of persons 70 years of age and older from the Supplement on Aging to the 1984 National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics. The purpose and design of this survey is described elsewhere.[11,12]

Of the persons selected for reinterview in 1986, 4270 met our criteria: they had had a follow-up interview or their vital status was known, and complete information was available for them on variables germane to this study.

Those who had died prior to follow-up and those who were missing pertinent follow-up information were deleted from disability outcome analyses; thus the sample size for this outcome was 3706.

The 1984 baseline interview included self-reported data on demographic variables, health conditions (including history of falls in the past year), and functional status; the 1986 follow-up reported vital status and functional status but did not obtain additional information about health conditions or falls.

Outcome Variables

Mortality. Mortality at follow-up was defined as death validated through a match with the National Death Index (NDI), or report of death for those for whom no NDI match was found. Five hundred twenty-five of the 4270 individuals used in this analysis had died during the 2-year follow-up period; cause of death was not available at the time of this analysis.

Disability. Functional disability at both baseline and follow-up was determined by reported difficulty with a modified selection of ADLs: bathing, dressing, eating, transferring from bed or chair, walking, going out, and toileting. Individuals were coded as having difficulty with an activity if they responded "yes" when asked if they had difficulty accomplishing that activity. Missing responses were deleted, as were those from individuals who reported that they did not perform the activity for reasons unrelated to health or physical ability. "Don't do" responses constituted less than 0.5% of responses for each individual ADL.

The dichotomous disability outcome variable was defined as having difficulty with a greater number of ADLs in 1986 than in 1984. Individuals having difficulty with fewer ADLs at follow-up (884, or 20.7%), as well as those whose functional ability had not changed, were defined as not having decreased functional ability.

Further multivariate analysis of the association of falls and subsequent disability was carried out with a categorical disability outcome variable: the number of ADLs reported at follow-up.

Exposure Variables

Respondents to the 1984 interview were asked if they had fallen in the past 12 months. If they answered "yes" they were asked if they had fallen once, or more than once; 401 individuals reported one fall and 405 reported more than one. Single and multiple fallers were compared separately to the reference group (those who had not fallen). In the multivariable analyses this comparison was accomplished with two dummy variables.

Covariates

Covariates used in the multivariable logistic regression models are listed in Table 1.

Because 14% of the baseline responses and 33% of the follow-up responses were completed by proxy respondents, and proxy status was significantly associated with multiple falls at baseline, a variable was included in each multivariable model to indicate proxy status of the respondent.

Other covariates were included in the models if they had been shown to be associated with falls in the LSOA population or in previous studies or (2) if they had been shown to be associated with mortality or disability in the LSOA population or in previous studies, and (3) if, in addition, the variable had sufficient responses for meaningful analysis.

Covariates were categorized as demographic, chronic condition, or disability variables. To examine the effect of specific covariates, we added all variables in a given category to the model as a block. Additionally, because our primary interest was to examine the effect of controlling for multiple chronic conditions and disabilities rather than to examine the effects of specific covariates upon the outcomes, we created two ordinal variables to represent the numbers of reported chronic conditions (0–11) and disabilities (0–7).

Demographic covariates were age in 1984, sex, marital status, and education (two dummy variables were used to describe "low" and "high" education, compared with the reference category of 9–12 years). Those missing data for education were assigned mean values (9.56 years for males, 9.99 for females). Education was included both as a surrogate for socioeconomic status (because 20% of the respon-

dents had not answered the question about income) and because there appears to be an inverse relation between educational attainment and mortality in middle-aged and older Americans.[13]

Chronic condition variables included osteoporosis/hip fracture, arthritis, hypertension, heart disease (myocardial infarction, rheumatic heart disease, coronary heart disease, and "other heart attack"), stroke, diabetes, vascular disease, cancer, visual deficits, hearing deficits, and thinness (body mass index < 21 kg/m^2). Low body mass index was included because the body weight at which minimal mortality occurs increases with advancing age[14] and the curve of mortality risk versus declining body mass index rises steeply for body mass indexes of 21 and less in middle-aged and older females.[15]

Disability covariates consisted of reported difficulty with individual ADLs at baseline (1984), as described above.

Analytical Methods

Our analytic strategy was to determine risks for 2-year mortality and functional decline in those with and without a history of falling, and then to examine the same questions in multivariable models, controlling for demographic characteristics, preexisting health conditions, and disabilities that could be related to an individual's propensity to fall as well as to risk of death or future disability.

Statistical analyses were performed with the SAS package.[16] All univariate and multivariate analyses were weighted; the results are thus applicable to the US population age 70 and over.

Multiple logistic regressions were performed with the LOGIST procedure.[17] The RTILOGIT procedure[18] was used with mortality and dichotomous disability outcome models to generate variances incorporating the weighting system used to compensate for the complex sampling design of the survey. RTILOGIT could not be used in models with an ordinal outcome variable; however, because the weighted variances generated in the dichotomous outcome model were so close to those obtained with the LOGIST procedure (within .0001 in all cases), we used standard deviations generated from the LOGIST procedure to calculate confidence intervals for the ordinal outcome models. Before ordinal outcome regressions were run, the partial proportional odds model[19] was used to test for nonproportionality of increasing response (number of ADLs) with exposure (falls); because this was nonsignificant, proportionality, or a common odds ratio between any dichotomous division in the outcome variable scale, could be assumed.

All multivariable models included the block of demographic covariates and a variable indicating the proxy status of the respondent.

Results

Descriptive and Cross-sectional Statistics

Unweighted descriptive statistics for the LSOA subpopulation used in this analysis are shown in Table 2. The extent to which the covariates used in the multivariable analyses are associated with having fallen is shown in Table 3. Compared with those who had not fallen, those who had fallen once in the year prior to baseline interview were significantly more likely to be female, age 80 or over, thin (BMI < 21), and unmarried, and to report the presence of visual or hearing difficulties, osteoporosis/hip fracture, arthritis, vascular disease, and difficulty with each of the ADLs. Those who had fallen more than once had characteristics similar to those listed above, and in addition were more likely to have less than 9 years of formal education, to have had the baseline questionnaire completed by a proxy, and to report the presence of hypertension, heart disease, stroke, diabetes, and cancer.

Mortality Outcome

Of those who reported no falls in the year prior to their 1984 interview, 11.9% were deceased by 1986, compared with 18% of those reporting one fall and 25.4% of those reporting two or more falls. The corresponding crude odds ratios were 1.5 (95% confidence interval [CI], 1.1–2.0) for single falls and 2.2 (95% CI, 1.7–2.8) for multiple falls (Table 4).

In the multivariable analysis, the odds ratios for single and multiple falls decreased slightly in magnitude, to 1.4 and 2.0, respectively, but remained significant at $P < .05$ when demographic variables and proxy status were controlled for (Table 5). When the block of chronic conditions was added to the model, the mortality odds ratios decreased further. With disability at baseline added to the model with demographic and proxy variables, the odds ratio for death decreased even more, losing significance for single falls, and when chronic conditions and disabilities were added together, the mortality odds ratios for single and multiple falls converged at 1.3 (95% CIs, 0.9–1.7 and 0.9–1.8, respectively). Substituting number of chronic conditions or disabilities for the individual conditions had almost no effect on these results.

Significant covariates in the full model (containing demographics, proxy status, individual chronic conditions, and individual disabilities) were age, male sex, high education (protective), thinness, arthritis (protective), stroke, cancer, difficulty bathing, and difficulty toileting. In the model with number of chronic condi-

TABLE 2—Characteristics of the Study Population

	n	%
Age in 1984, y		
70–74	1497	35.1
75–79	1082	25.3
80–84	1040	24.4
85+	651	15.2
Sex		
Female	2715	63.6
Male	1555	36.4
Race		
White	3799	88.8
Non-White	471	11.2
Marital status		
Unmarried	2286	53.5
Married	1984	46.5
Education, y		
0–8	1733	40.6
9–12	1788	41.9
13+	749	17.5
Body mass index (kg/m^2)		
<21	843	19.7
21–29	2986	69.9
30+	441	10.3
Falls in year prior to 1984		
None	3289	77.0
One	473	11.1
Two or more	508	11.9
Chronic conditions in 1984		
None	338	7.9
1–2	1782	41.7
3–4	1604	37.6
5+	546	12.8
ADL difficulties in 1984		
None	3111	72.9
1–2	682	16.0
3–4	263	6.2
5+	214	5.0
ADL difficulties in 1986		
None	2226	60.1
1–2	828	22.3
3–4	325	8.8
5+	327	8.8
Vital status in 1986		
Alive	3745	87.7
Deceased	525	12.3
Change in number of ADL difficulties 1984–1986		
Decrease	884	20.7
No change	2251	52.7
Increase	1135	26.6

Note. ADLs = activities of daily living.

TABLE 3—Weighted Cross-sectional Associations of Demographic Variables, Chronic Conditions, and Activities of Daily Living with Falls

	OR, One Fall vs No Falls (95% CI) (n = 3762)	OR, Two or More Falls vs No Falls (95% CI) (n = 3797)
Demographics		
Age (80+ vs 70–79)	1.7 (1.4–2.1)	2.2 (1.8–2.6)
Sex (male vs female)	0.6 (0.5–0.8)	0.7 (0.6–0.9)
Marital status (Married vs unmarried)	0.6 (0.5–0.7)	0.6 (0.5–0.8)
Education		
(<9 vs ≥9 years)	1.1 (0.9–1.3)	1.4 (1.2–1.7)
(>12 vs ≤12 years)	1.1 (0.8–1.4)	0.9 (0.7–1.2)
Chronic Conditions		
Body mass index (<21 vs ≥21)	1.5 (1.2–1.9)	1.4 (1.1–1.8)
Visual deficit	1.5 (1.2–1.8)	2.4 (1.9–2.8)
Hearing deficit	1.3 (1.1–1.6)	2.0 (1.7–2.4)
Osteoporosis/hip fracture	1.8 (1.3–2.5)	1.6 (1.1–2.2)
Arthritis	1.6 (1.3–1.9)	2.4 (1.9–2.9)
Hypertension	1.2 (1.0–1.5)	1.6 (1.3–1.9)
Heart disease	1.0 (0.8–1.3)	2.3 (1.8–2.8)
Stroke	1.4 (1.0–2.1)	3.6 (2.8–4.7)
Vascular disease	1.5 (1.1–1.9)	2.0 (1.5–2.6)
Diabetes	1.1 (0.8–1.5)	2.1 (1.6–2.7)
Cancer	1.1 (0.8–1.5)	1.9 (1.5–2.5)
Disabilities		
Difficulty bathing	2.2 (1.7–2.9)	4.9 (3.9–6.1)
Difficulty dressing	2.0 (1.4–2.8)	4.7 (3.6–6.1)
Difficulty eating	3.5 (1.8–6.7)	9.8 (6.4–15.0)
Difficulty walking	2.1 (1.6–2.6)	4.8 (3.9–5.8)
Difficulty transferring	2.3 (1.7–3.1)	4.5 (3.5–5.7)
Difficulty going out	2.1 (1.6–2.8)	5.2 (4.2–6.4)
Difficulty toileting	2.1 (1.4–3.2)	4.5 (3.3–6.1)
Self vs proxy	0.9 (0.6–1.2)	1.4 (1.1–1.8)

Note. OR = odds ratio; CI = confidence interval.

TABLE 4—Mortality Rates and Crude Mortality Odds Ratios

Falls in 1984	Alive in 1986	Deceased in 1986 (%)	Odds Ratio (95% CI)
None	2939	350 (11.9)	—
One (n = 3762)	401	72 (18.0)	1.5 (1.1–2.0)
Two or more (n = 3797)	405	103 (25.4)	2.2 (1.7–2.8)

Note. CI = confidence interval.

tions and disabilities, the same demographic covariates were significant, along with the number of chronic conditions and number of disabilities.

Disability Outcome

In the univariate analysis (Table 6), single falls were not significantly associated with subsequent increased disability, but those who had fallen more than once were more than twice as likely to report difficulty with additional ADLs as were those who had fallen once.

Single falls did not predict increased disability in any of the multivariate models, whereas multiple falls were significant predictors in all models, both dichotomous and ordinal (Tables 7 and 8). As with the mortality outcome models, the addition of successive groups of covariates caused a decrease in the magnitude of the disability odds ratios for multiple falls. Substitution of numbers of chronic conditions or ADL difficulties present at baseline for the groups of individual conditions had little effect on the odds ratios.

A similar trend was seen in the categorical outcome models, with the disability odds ratio for multiple falls decreasing in magnitude but remaining significantly greater than 1.0 as successive blocks of covariates were added (Table 8).

Significant covariates in the categorical full model with individual chronic conditions were age, being married (protective), osteoporosis/hip fracture, arthritis, hypertension, heart disease, stroke, diabetes, proxy status, and number of functional disabilities at baseline. In the model with number of chronic conditions rather than individual chronic conditions, there was a significant association between this variable and decreased functional ability.

Discussion

Although the univariate analysis demonstrated a significant association between falls and 2-year mortality in this nationally representative population, the results of the multivariable analyses support the hypothesis that elevated mortality associated with falling in older persons is largely due to the presence of multiple chronic diseases and disabilities in those who are prone to fall. Falling is not an independent predictor of mortality when chronic disease and functional disability are controlled for, but it tends to be associated with these conditions in our baseline data. Our findings are consistent with other studies implicating underlying disease and disability as possible causal factors in falling.

Single falls do not appear to be associated with increased functional disability over a 2-year follow-up, but multiple falls are, and they remain significant, if not independent, predictors even when chronic conditions and functional disabilities present at baseline are controlled for. This could be due to injuries not included in the model, to other chronic diseases associated with both falls and disability that we were not able to control for (e.g., parkinsonism, Alzheimer's disease), or to these and other chronic diseases that might be undiagnosed and thus unreported. It has also been noted that fear of falling can cause people to restrict their normal activities[8]; it is possible that a repeat faller might limit his or her activities to the point where ADLs are affected. Although a relatively small proportion of falls result in serious injury (such as fractures or severe soft tissue trauma), it is possible that less serious injuries (as well as outcomes such as anxiety about further falls and loss of confidence) may disproportionately affect older persons.

In light of Tinetti's findings implicating number of conditions as one of the most important risk factors for falls, it is

TABLE 5—Odds Ratios (and 95% Confidence Intervals) from Multiple Logistic Regression Analysis: Falls as Predictors of 2-Year Mortality, when Demographic (DM), Chronic Condition (CC), and Disability (DB) Covariates Are Controlled For (n = 4270)

	Covariates						
Falls in 1984	DM	DM, CC	DM, DB	DM, CC, DB	DM, No. CC	DM, No. DB	DM, No. CC, No. DB
One	1.4 (1.1–1.9)	1.4 (1.0–1.9)	1.3 (0.9–1.8)	1.3 (0.9–1.8)	1.3 (1.0–1.8)	1.3 (0.9–1.7)	1.2 (0.9–1.7)
Two or more	2.0 (1.5–2.6)	1.6 (1.2–2.1)	1.4 (1.1–1.9)	1.3 (0.9–1.7)	1.6 (1.2–2.1)	1.4 (1.1–1.9)	1.3 (1.0–1.7)

Note. All models are weighted and control for proxy responses.

interesting that controlling for number of chronic diseases or functional disabilities present at baseline had an effect on the mortality and disability odds ratios for falls almost identical to the effect of controlling for the individual conditions. An examination of this LSOA subpopulation suggests a possible explanation with regard to chronic condition variables in the mortality outcome models. As shown in Table 1, the majority of individuals (nearly 80%) reported having from one to four of the chronic conditions we examined in our models. The most prevalent chronic conditions reported were arthritis (55.2%), hypertension (46.1%), hearing deficits (39.7%), and visual deficits (37.0%), none of which was a significant risk factor for mortality in the full model. An assumption that the majority of people reporting chronic conditions at baseline suffer from more than one of these prevalent but nonfatal conditions could explain the similar effect of controlling for number of conditions as opposed to individual conditions. In support of this assumption, a recent analysis of comorbidities in respondents aged 60 and older to the National Health Interview Survey and the Supplement on Aging found that arthritis and hypertension were the most common comorbid conditions, reported by 24.1% of that population, followed by arthritis and cataract (11.7%).[20]

These data provide an interesting note on the relationship between single and multiple falls. Although the crude mortality odds ratio was much higher for multiple than for single falls, the difference between the two decreased as covariates representing "intrinsic" characteristics were added to the model. When demographic, chronic condition, and disability variables were controlled for simultaneously, the odds ratios and test-based confidence intervals were nearly identical for single and multiple falls, and neither were significant predictors of mortality. On the other hand, multiple falls predicted increased disability and single falls did not, even when all covariates were added to the model. The most likely source of bias in comparisons of single and multiple falls—that a substantial number of single-fallers had additional falls prior to follow-up—would tend to obscure this difference. Perhaps there is a temporal difference; those experiencing multiple falls earlier in time may be more likely to become disabled sooner.

TABLE 6—Functional Ability Status and Crude Disability Odds Ratios

Falls in 1984	No Additional Disability in 1986[a]	Additional Disability in 1986	Odds Ratio (95% CI)
None	2090	819	—
One (n = 3306)	271	126	1.2 (0.9–1.5)
Two or more (n = 3309)	210	190	2.2 (1.8–2.7)

Note. CI = confidence interval.
[a]Includes those reporting difficulty with fewer activities of daily living (ADLs) in 1986 than in 1984.

There are also implications for caregivers and clinicians in the interrelationships of mortality, disability, chronic conditions, and falls. As individuals accumulate larger numbers of chronic conditions and disabilities—even those that are individually relatively benign—their potential for experiencing adverse outcomes from falls increases substantially. These individuals may be most in need of interventions to reduce both the risk and the severity of falls. Likewise, multiple falls in older persons should be taken seriously by physicians and other caregivers as indicators of possible underlying conditions that could increase risk for functional dependency or death within the next few years. □

Acknowledgment

This work was supported by a grant from the National Institute on Aging (AG 06996-02).

References

1. Prudham D, Grimley Evans J. Factors associated with falls in the elderly: a community study. *Age Aging.* 1981;10:141–146.
2. Campbell AJ, Reinken J, Allen BC, Martinez GS. Falls in old age: a study of frequency and related clinical factors. *Age Aging.* 1981;10:264–270.
3. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. *N Engl J Med.* 1988;319:1701–1707.
4. Perry BC. Falls among the elderly. *J Am Geriatr Soc.* 1982;30:367–371.
5. Champion HR, Copes W, Buyer D, Flanagan ME, Bain L, Sacco WJ. Major trauma in geriatric patients. *Am J Public Health.* 1989;79:1278–1282.
6. Sattin RW, Lambert Huber DA, DeVito CA, et al. The incidence of fall injury events among the elderly in a defined population. *Am J Epidemiol.* 1990;131:1028–1037.
7. Wild D, Nayak USL, Issacs B. How dangerous are falls in old people at home? *Br Med J.* 1981;282:266–268.
8. Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls. *JAMA.* 1989;261:2663–2668.
9. Nickens H. Intrinsic factors in falling among the elderly. *Arch Intern Med.* 1985;145:1089–1093.
10. Tinetti ME, Williams TF, Mayewski R. Fall risk index for elderly patients based on number of chronic disabilities. *Am J Med.* 1986;80:429–434.
11. Fitti JE, Kovar MG. *The Supplement on Aging to the 1984 National Health Interview Survey.* Washington, DC: National Center for Health Statistics; 1987. Vital and

TABLE 7—Odds Ratios (and 95% Confidence Intervals) from Multiple Logistic Regression Analysis: Falls as Predictors of Additional Disability, Dichotomous Outcome, when Demographic (DM), Chronic Condition (CC), and Disability (DB) Covariates Are Controlled For (n = 3706)

	Covariates						
Falls in 1984	DM	DM, CC	DM, DB	DM, CC, DB	DM, No. CC	DM, No. DB	DM, No. CC, No. DB
One	1.1 (0.9–1.4)	1.0 (0.8–1.3)	1.1 (0.9–1.4)	1.0 (0.8–1.3)	1.0 (0.8–1.3)	1.1 (0.9–1.4)	1.0 (0.8–1.3)
Two or more	1.9 (1.5–2.3)	1.4 (1.1–1.8)	1.9 (1.5–2.4)	1.6 (1.2–2.0)	1.5 (1.2–1.9)	2.0 (1.5–2.5)	1.6 (1.3–2.0)

Note. All models control for proxy responses.

TABLE 8—Odds Ratios (and 95% Confidence Intervals) from Ordinal Outcome Multiple Logistic Regression Analysis: Falls as Predictors of Increase of One Difficult Activity of Daily Living Between 1984 and 1986, when Demographic (DM) and Chronic Condition (CC) Covariates Are Controlled For (n = 3706)

	Covariates		
Falls in 1984	DM	DM, CC	DM, No. CC
One	1.1 (0.9–1.4)	1.0 (0.8–1.3)	1.0 (0.8–1.3)
Two or more	2.1 (1.7–2.5)	1.7 (1.4–2.1)	1.8 (1.4–2.2)

Note. All models control for proxy responses and number of activities of daily living (ADLs) at baseline.

Health Stat. 1(21). US Dept of Health and Human Services publication PHS 87-1323.

12. Kovar MG. *The Longitudinal Study of Aging: The 1986 Reinterview Public Use File.* Hyattsville, Md: US Dept of Health and Human Services; 1987: 45–46. Proceedings of the 1987 Public Health Conference on Records and Statistics. US Dept of Health and Human Services publication PHS 88-1214.
13. Feldman JJ, Makuc DM, Kleinman JC, Cornoni-Huntley J. National trends in educational differentials in mortality. *Am J Epidemiol.* 1989;129:919–933.
14. Andres R, Elahi D, Tobin JD, Muller DC, Brant L. Impact of age on weight goals. *Ann Int Med.* 1985;103:1030–1033.
15. Waaler HT. Height, weight, and mortality: the Norwegian experience. *Acta Med Scand.* 1984;679(suppl):1–56.
16. SAS Institute Inc. *SAS User's Guide: Basics Version 5 Edition.* Cary, NC: SAS Institute Inc; 1985.
17. Harrell FE Jr. The LOGIST procedure. In: *SUGI Supplemental Library User's Guide.* Cary, NC: SAS Institute Inc; 1986:269–294.
18. Shah BV, Folsom RE, Harrell FE, Dillard CN. *RTILOGIT: Procedure for Logistic Regression on Survey Data.* Research Triangle Park, NC: Research Triangle Institute; 1987.
19. Peterson B, Harrell FE Jr. Partial proportional odds models for ordinal response variables. *Appl Stat.* 1990;39:205–217.
20. Guralnik JM, LaCroix AZ, Everett DF, Kovar MG. Aging in the eighties: the prevalence of comorbidity and its association with disability. *Adv Data.* 1989;170:1–8.

Living Arrangements and Survival among Middle-Aged and Older Adults in the NHANES I Epidemiologic Follow-up Study

ABSTRACT

Background. There is concern about but little information on how living alone affects the health and survival of older adults.

Methods. We examined the association between living arrangements (living alone, with a spouse, or with someone other than a spouse) and survival among 7651 adults, aged 45 to 74 years in the National Health and Nutrition Examination Survey (NHANES I) (1971–1975) and traced at the NHANES I Follow-up Study (1982–1984), to see whether certain sociodemographic factors (race, education, income, and employment), health behaviors (alcohol, smoking, physical activity, and obesity), or chronic medical conditions were influential in the association.

Results. We found a stronger association of living arrangements with survival for men than for women, and for middle-aged men than for older men. For men, those living alone and those living with someone other than a spouse were equally disadvantaged in terms of survival. Income, race, employment, and physical activity influenced the association of living arrangements and survival, but their impact varied by age, gender, and living arrangement.

Conclusion. Living arrangements had a weak impact on survival among men, but had no effect among women. (*Am J Public Health.* 1992;82:401–406)

Maradee A. Davis, PhD, MPH, John M. Neuhaus, PhD, Deborah J. Moritz, PhD, and Mark R. Segal, PhD

Introduction

In the United States, the number and proportion of elderly persons who live alone have increased dramatically since 1960.[1] As of 1989, approximately 30% of all persons aged 65 years and older lived alone, and almost four fifths of those solitary elders were women.[2] There has been a trend for older adults to live alone if they are no longer married, and this pattern is likely to continue.[3]

Numerous studies indicate that married persons have lower mortality rates than unmarried ones[4–13] and that differences in mortality by marital status are more pronounced for men than for women.[9–13] However, few studies have addressed whether there is additional mortality risk associated with being unmarried and living alone, or whether living arrangements affect survival more or less strongly for older adults compared with middle-aged adults.

There is concern that older adults who live alone may be particularly vulnerable to poverty, social isolation, loneliness, decreased psychological well-being, and adverse health outcomes.[14–17] And there is considerable evidence among older or middle-aged adults that mortality is associated with sociodemographic characteristics[18–24] (age, sex, race, education, income, and employment), health behaviors[19,22,25–27] (alcohol consumption, smoking, obesity, and physical activity), and health status.[19,28,29] However, there is little existing epidemiological information about the quality of life, health behaviors, and health status of the elderly who live alone, or about the impact of living alone on survival.[30]

Because many older adults want to live independently, they may be likely to live alone until their health declines to the point where they can no longer function independently; they then move in with others or are institutionalized.[31,32] Thus, those living with someone other than a spouse may be at higher risk of mortality because their poor health led them to live with someone.

Evidence from four national data sets suggests that age, sex, socioeconomic status, and other factors may influence the diet and health of the elderly living alone.[30,33–35] These findings highlight the need to examine whether specific sociodemographic, behavioral, or health characteristics that may differ by type of living arrangement affect the association of living arrangements with survival. In this report, we examine the association of living arrangements with subsequent survival over a 7- to 13-year period for a sample of US adults aged 45 to 74 years from the National Health and Nutrition Examination Survey 1971–1975 (NHANES I)[36,37] who were traced in the NHANES I Epidemiologic Follow-up Study 1982–1984 (NHEFS).[38] We focused on four research questions: (1) Do persons living alone have shorter survival times than those living with a spouse or with someone other than a spouse? (2) Are living arrangements more strongly associated with survival time for men

The authors are with the Department of Epidemiology and Biostatistics, School of Medicine, University of California, in San Francisco.

Requests for reprints should be sent to Maradee A. Davis, PhD, MPH, Department of Epidemiology and Biostatistics, University of California, San Francisco, CA 94143–0560.

This paper was submitted to the Journal January 31, 1991, and accepted with revisions June 19, 1991.

TABLE 1—Sample Sizes and Number of Deaths (NHEFS 1982–1984), by Age, Sex, and Living Arrangements (NHANES I)

	Living Arrangements					
	Alone		With Other Than Spouse		With Spouse	
	Sample	Deaths	Sample	Deaths	Sample	Deaths
Men						
45–54 years	65	15	74	17	872	92
55–64 years	71	23	69	25	695	135
65–74 years	214	117	155	89	1400	629
Women						
45–54 years	91	7	185	16	889	38
55–64 years	160	18	135	16	635	66
65–74 years	644	203	391	123	906	253

than for women? (3) Are living arrangements more strongly associated with survival time for older adults than for middle-aged adults? (4) Do specific sociodemographic characteristics, health behaviors, or chronic medical conditions known to be associated with mortality account for patterns in which living arrangements are associated with survival?

Methods

Study Sample

The NHANES I, a national probability sample of the civilian, noninstitutional population aged 1 to 74 years, collected demographic, medical history, clinical, and laboratory information.[36,37] The NHEFS 1982–1984 attempted to trace all 14 407 NHANES I persons aged 25 to 74 years to determine whether they were still alive and to conduct in-depth interviews, collect selected physical measurements, and obtain health facility records and death certificates.[38]

Analyses in this report are restricted to the 7957 persons aged 45 to 74 years at NHANES I, of whom 7651 (96%) were traced at follow-up. Among those traced, 5769 (75%) were alive and 1882 (25%) had died.[38] Death certificates were obtained for 96% of the deaths, and proxy reports of death were obtained for the remainder.[38]

Mean length of follow-up for the survivors was 10 years. Nonsurvivors died an average of 5.5 years after NHANES I. Madans and colleagues have reported that the survival experience of the 1982–1984 NHEFS cohort is representative of the survival experience of the total US population.[39] Table 1 describes the sample sizes and number of deaths by age, sex, and living arrangements for persons included in these analyses.

Variables

Baseline (NHANES I) living arrangements were categorized as living alone, living with a spouse, or living with someone other than a spouse. Baseline sociodemographic variables included age (45 to 54 years, 55 to 64 years, 65 to 74 years), sex, race (White or other), education, income (dollars per household member), and employment (employed or not, employed outside the home, during the previous 3 months). Those not employed included those who were keeping house, retired, laid off from work, looking for work, or ill or unable to work during the previous 3 months. Persons with missing income data (n = 319) were excluded from analyses involving income.

Health behavior variables included body mass index (weight [kg]/height [m^2]), physical activity, alcohol consumption, and smoking. The measures of physical activity were nonrecreational activity (inactive, moderately active, or very active) and recreational activity (little or none, moderate, much). Daily alcohol consumption was divided into four categories: none, one drink per day, two drinks per day, or more than two drinks per day, based on questions about the subject's usual 24-hour frequency of alcohol consumption. Smoking behavior was categorized as having never smoked, having formerly smoked, or currently smoking. Only a random subsample of the NHANES I sample was asked about smoking (n = 3968 for those in this report), whereas all respondents or proxies of decedents were asked about smoking at the follow-up. For those persons not asked about smoking at baseline (n = 3080 for these analyses), baseline smoking status was determined from a lifetime smoking history ascertained at NHEFS 1982–1984. A high level of agreement has been found between NHANES I baseline smoking responses and baseline smoking as determined in the NHEFS.[40] For persons who were missing smoking information at both NHANES I and NHEFS (n = 603 for these analyses), smoking status was imputed using the method developed by Kleinman and colleagues.[41]

History of chronic medical conditions was ascertained from questions about whether a doctor had ever told the respondents that they had any of 36 conditions. In addition, blood pressure readings were obtained, and persons were classified as hypertensive if they had either a systolic reading of 165 mm Hg or greater or a diastolic reading of 95 mm Hg or greater, or were taking medication for hypertension. For these analyses, 19 of the conditions that are frequent causes of death[42] were grouped into nine categories (pulmonary disease, heart disease, hypertension, stroke, gastrointestinal disease, diabetes, kidney disease, malignant tumor, and fracture). The number of reported disease categories was counted and categorized (none, one, or two or more categories).

Statistical Analyses

We examined the association of living arrangements from NHANES I with subsequent survival from NHEFS using Cox proportional hazards models,[43] which were calculated by the PHGLM procedure in SAS.[44] Separate proportional hazards models were fitted in each of six NHANES I age-sex groups. Using a standard graphical method,[43] we examined the adequacy of the proportional hazards assumption by plotting the log-minus-log survivor function by living arrangement group within each of the six age-sex groups. For each age-sex group, the plots exhibited nearly constant separation, suggesting that the assumption of proportional hazards was reasonable.

We initially computed estimates of the age-adjusted relative hazards corresponding to living arrangements, and then we examined the role of sociodemographic characteristics, health behaviors, and chronic conditions in the living arrangements–survival association by calculating the relative hazards estimates adjusted for these variables. All analyses were also adjusted for age (in individual years) within each age group. Interactions between living arrangements and each of the sociodemographic, behavioral, and chronic conditions variables were examined by including products of the indicator

variables for living arrangements and each variable as covariates in the proportional hazards models and testing the statistical significance of the interaction parameters using the likelihood ratio test. No interactions were found.

The NHANES I sampling design, which included differential sampling rates by age, sex, and income, was incorporated into our analyses. The sample proportions of persons in the three living arrangement groups were calculated using the NHANES I sample weights (Figure 1). Additionally, all proportional hazards models were fitted separately by sex- and design-compatible age groups while the final models were adjusted for income and race. Our proportional hazards analyses did not incorporate the NHANES I cluster sample design into the standard error estimates. Similar analyses of NHANES I survival data found design effects for Cox regression coefficients near 1.[20]

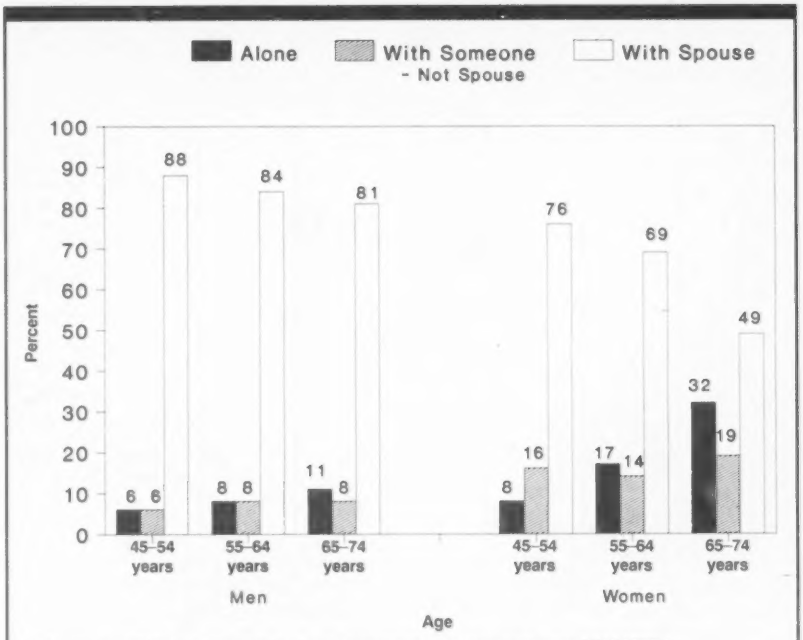


Figure 1—Distribution of living arrangements by age and sex: NHANES I.

Results

Distribution of Living Arrangements

The percentage of persons married and living with a spouse decreased with increasing age (Figure 1), but the decrease was greater for women than for men. Among men who were not married and living with a spouse, approximately equal proportions lived alone as lived with someone other than a spouse. For women, the proportion living alone increased with age.

Association of Living Arrangements with Survival

Men who lived alone or with someone other than a spouse had significantly shorter survival times compared with those living with a spouse (Table 2). The magnitude of the relative hazard of dying was similar within each age group for men living alone and men living with someone other than a spouse. The age-adjusted relative hazard of dying was highest in the youngest age group, 45 to 54 years, and decreased somewhat with increasing age (Table 2). For women, there was a weak association of living arrangements with survival only for those aged 45 to 54 years old (Table 2).

The Role of Chronic Medical Conditions, Health Behaviors, and Sociodemographic Characteristics

Chronic conditions. The presence of chronic conditions at NHANES I decreased survival time in each age-sex group (data not shown). However, adjusting for the number of chronic conditions did not reduce the association of living arrangements with survival (Table 2); this is likely due to the lack of association of chronic conditions with living arrangements (data not shown). Nor was the association of living arrangements with survival reduced when we adjusted individually for each of the nine categories of chronic conditions (data not shown).

Health behaviors. Behavioral variables did not appear to have a major influence on the association of living arrangement with survival, except among men aged 55 to 64 years who were living alone (Table 2). Among the behavioral variables examined, physical activity resulted in the largest reduction in the relative hazard. Adjusting additionally for number of chronic conditions did not change the magnitude of the relative hazards. We also examined the role of self-reported recreational activity in accounting for the association of living arrangements with survival. The relative hazards adjusting for recreational activity (data not shown) were similar to those in Table 2 adjusting for nonrecreational activity.

Sociodemographic characteristics. For men, adjusting for race, education, and income individually did not substantially reduce the association of living arrangements with survival (Table 2). However, in the 45- to 54-year-old group, adjusting for employment substantially reduced the relative hazards when both men living alone and those living with someone other than a spouse were compared with those living with a spouse (Table 2). Adjusting additionally for number of chronic conditions produced relative hazards of death that were similar to those adjusted solely for employment and age (relative hazard [RH] 1.9, confidence interval [CI] 1.1, 3.4 for those living alone; and RH 1.6, CI 0.9, 2.7 for those living with others). With simultaneous adjustment for all variables, living arrangements among men were still marginally associated with survival.

For women aged 45 to 54 years, adjusting for race and income reduced the relative hazard for those living with someone other than a spouse compared with those living with a spouse (Table 2), whereas adjusting for education or employment did not. Adjusting simultaneously for race, income, age, and number of chronic conditions produced relative hazards similar to those produced when adjusting only for age and race or age and income (data not shown). Because of concern about the impact of living arrangements on older, low-income women, we examined the impact of living arrangements on survival for five different income strata in women aged 65 to 74 years. The magnitude of the relative hazards for both women living alone and those living with

TABLE 2—Relative Hazards of Death (NHEFS 1982–1984) for Men and Women Living Alone or with Someone Other Than a Spouse Compared with Those Living with a Spouse: Adjusted for Chronic Conditions, Health Behaviors, and Socioeconomic Characteristics (NHANES I)

	Aged 45–54 Years				Aged 55–64 Years				Aged 65–74 Years			
	Alone		With Others		Alone		With Others		Alone		With Others	
	RH	95% CI	RH	95% CI	RH	95% CI	RH	95% CI	RH	95% CI	RH	95% CI
Men												
Age adjusted	2.4	(1.4, 4.1)	2.4	(1.4, 4.0)	1.8	(1.2, 2.8)	2.0	(1.3, 3.1)	1.3	(1.1, 1.5)	1.4	(1.1, 1.7)
Adjusted for number of chronic conditions:[a]												
Chronic conditions	2.3	(1.3, 3.9)	2.4	(1.4, 4.0)	1.7	(1.1, 2.7)	2.2	(1.4, 3.4)	1.3	(1.1, 1.6)	1.3	(1.0, 1.6)
Adjusted for health behavior variables:												
Alcohol	2.2	(1.3, 3.8)	2.3	(1.4, 3.9)	1.8	(1.1, 2.7)	1.9	(1.2, 2.9)	1.3	(1.0, 1.6)	1.4	(1.1, 1.7)
Smoking	2.3	(1.3, 3.9)	2.1	(1.3, 3.6)	1.7	(1.1, 2.6)	2.0	(1.3, 3.1)	1.2	(1.0, 1.5)	1.3	(1.0, 1.5)
BMI	2.3	(1.4, 4.1)	2.3	(1.4, 4.0)	1.7	(1.1, 2.7)	1.9	(1.2, 2.9)	1.2	(1.0, 1.5)	1.3	(1.0, 1.6)
Activity	2.3	(1.3, 4.0)	2.3	(1.4, 3.8)	1.5	(1.0, 2.4)	2.0	(1.3, 3.1)	1.3	(1.0, 1.5)	1.2	(1.0, 1.5)
Adjusted for socioeconomic variables:												
Race	2.1	(1.2, 3.6)	2.2	(1.3, 3.7)	1.7	(1.1, 2.7)	2.0	(1.3, 3.0)	1.2	(1.0, 1.5)	1.3	(1.1, 1.7)
Education	2.3	(1.3, 3.9)	2.1	(1.2, 3.6)	1.8	(1.1, 2.7)	1.9	(1.2, 3.0)	1.2	(1.0, 1.5)	1.2	(1.0, 1.6)
Income	2.8	(1.6, 4.9)	2.2	(1.3, 3.7)	2.0	(1.3, 3.2)	1.8	(1.1, 2.8)	1.3	(1.1, 1.6)	1.3	(1.1, 1.6)
Employment	1.9	(1.1, 3.2)	1.6	(0.9, 2.7)	1.7	(1.1, 2.6)	1.9	(1.2, 2.9)	1.2	(1.0, 1.5)	1.3	(1.0, 1.6)
Adjusted for all variables	1.7	(0.9, 3.1)	1.3	(0.7, 2.3)	1.5	(0.9, 2.4)	1.6	(1.0, 2.6)	1.2	(1.0, 1.5)	1.1	(0.8, 1.4)
Women												
Age adjusted	1.7	(0.8, 3.9)	2.0	(1.1, 3.7)	1.0	(0.6, 1.8)	1.2	(0.7, 2.0)	1.1	(0.9, 1.3)	1.1	(0.9, 1.3)
Adjusted for number of chronic conditions:[a]												
Chronic conditions	1.7	(0.8, 3.8)	2.0	(1.1, 3.5)	1.1	(0.6, 1.8)	1.1	(0.7, 2.0)	1.1	(0.9, 1.3)	1.0	(0.8, 1.3)
Adjusted for health behavior variables:[a]												
Alcohol	1.7	(0.7, 3.7)	2.0	(1.1, 3.7)	1.0	(0.6, 1.8)	1.3	(0.7, 2.2)	1.1	(0.9, 1.3)	1.1	(0.9, 1.3)
Smoking	1.7	(0.7, 3.7)	1.9	(1.1, 3.5)	1.0	(0.6, 1.7)	1.2	(0.7, 2.1)	1.1	(0.9, 1.3)	1.1	(0.9, 1.3)
BMI	1.6	(0.7, 5.0)	2.0	(1.1, 3.6)	1.0	(0.6, 1.8)	1.1	(0.7, 2.0)	1.1	(0.9, 1.3)	1.0	(0.8, 1.3)
Activity	1.7	(0.8, 3.9)	2.0	(1.1, 3.6)	1.0	(0.6, 1.7)	1.1	(0.6, 1.9)	1.0	(0.9, 1.3)	1.0	(0.8, 1.3)
Adjusted for socioeconomic variables:[a]												
Race	1.5	(0.7, 3.4)	1.5	(0.8, 2.8)	1.0	(0.6, 1.8)	1.2	(0.7, 2.0)	1.1	(0.9, 1.3)	1.0	(0.8, 1.3)
Education	1.7	(0.8, 3.9)	1.9	(1.1, 3.5)	1.0	(0.6, 1.8)	1.1	(0.7, 2.0)	1.1	(0.9, 1.3)	1.0	(0.8, 1.3)
Income	2.1	(0.9, 4.8)	1.4	(0.7, 2.5)	1.1	(0.6, 1.8)	1.1	(0.7, 2.0)	1.0	(0.9, 1.3)	1.0	(0.8, 1.2)
Employment	2.0	(0.9, 4.6)	2.2	(1.2, 4.0)	1.0	(0.6, 1.7)	1.1	(0.7, 2.0)	1.1	(0.9, 1.3)	1.1	(0.9, 1.4)
Adjusted for all variables	1.6	(0.7, 3.7)	1.1	(0.6, 2.0)	0.9	(0.5, 1.6)	1.2	(0.6, 2.1)	1.0	(0.8, 1.2)	1.0	(0.8, 1.2)

[a] All analyses controlled for age within each age group.
Note. RH = relative hazard; CI = confidence interval.

others was very similar across the income strata, and the interaction between income and living arrangements was statistically nonsignificant ($P = .60$).

Discussion

Our analyses were predicated on concern about the effects of living alone on the health and mortality of older adults. Although we did not find that the many women aged 65 to 74 years who lived alone were at higher risk of death than were women their age who were married—this finding being consistent with evidence from 1966–1968 US death and census data[5]—our analyses were restricted to women under 75 years of age. Among women over age 75, the impact of living alone on survival may be different. It is important to study persons aged 74 years and older because 40% of them live alone.[1]

There is growing evidence to suggest that older women living alone are not disadvantaged in terms of their social contacts,[30] mental health,[31,45] physical and functional health,[30,31] or dietary status.[33-35] Our findings also suggest that they are not disadvantaged in terms of survival. In addition, we did not find that older, low-income women living alone had poorer survival than higher-income women living alone. However, we used a measure of current income in these analyses and did not have measures of lifelong income, which may be more important for survival.

We did not find any evidence to suggest that living arrangements have an additional influence on mortality beyond marital status. Men aged 65 to 74 years who lived alone were at slightly higher risk of death than those who were married, as were those who lived with someone other than a spouse. The higher risk of mortality for older men living alone is consistent with results from previous studies of US census data,[5] Swedish men,[46] and widowed men in Washington County, Maryland.[13]

It is somewhat surprising that living with others did not offer a protective effect compared with living alone, given the results of various studies suggesting the importance of social support for decreased mortality.[21,47,48] However, it is possible that any advantage in this group that

comes from social support may be offset by the poor health that requires living with someone. We did not find that persons living with a nonspouse had more chronic conditions than those living alone, but there may be some differences in other aspects of health or functioning that may be related to increased mortality for this group.

We observed a somewhat stronger association of living arrangements with survival for middle-aged adults than for older adults, particularly for men, suggesting the possibility that the importance of living arrangements as a risk factor for mortality decreases with age. Attenuations in the magnitude of the associations of risk factors with survival in older age have been reported for marital status[5,21] as well as for other mortality risk factors.[21,22,49] It is possible that this attenuation is due to selective survival of the fittest as one ages.[21] It is also quite likely that the impact of living arrangements on health and mortality changes across the adult life cycle. For example, using data from the Alameda County Study, Seeman and colleagues[8] report that marital status assumed primary importance for those younger than 60 years old; however, ties with close friends and/or relatives assumed greater importance for those aged 60 years and older.

Although sociodemographic characteristics and health behaviors are known to affect mortality,[19,20,22] we observed only four of the variables (employment status, physical activity, race, income) to account, at least in part, for the association of living arrangements with survival, and then only in certain age and sex subgroups. Employment status was the primary variable associated with both living arrangements and survival for men aged 45 to 54 years. Those who were not working at baseline were at a substantially higher mortality risk than those who were working. We did not find that poorer health status in terms of chronic medical conditions accounted for this higher risk. However, it is possible that some other measures of health status, social support,[50] mental health, or behavioral characteristics not examined in this study may confound this association. In analyses not shown here, excluding men who died in the first 2 years of follow-up, the relationship between living arrangements, employment status, and survival did not change, indicating that the association of employment status with survival time was not due to imminent death in those men aged 45 to 54 years who were unemployed.

Previous studies have reported that controlling for baseline health status does not explain the higher mortality of those living alone[46] or unmarried.[51] Likewise, controlling for the number of chronic medical conditions present at NHANES I did not account for our observed association of living arrangements with mortality. However, the measures available to define health status on the full NHANES I sample were limited and did not include measures of duration or severity of the chronic conditions or physical functioning. More comprehensive measures of health status were available in the smaller, detailed component of the sample, but the small number of deaths in this subsample precluded meaningful analysis.

Interim changes in living arrangements between NHANES I and the follow-up may have obscured existing relationships between living arrangements and survival. However, the survival curves for living arrangements within age and sex groups conformed to the proportional hazards assumption throughout the follow-up period, including the early period when living arrangements were least likely to have changed.

By imputing survival times using several different imputation schemes, we investigated the possibility of nonresponse bias resulting from the exclusion of persons not traced in the NHEFS 1982–1984. Extreme assumptions regarding the survival time of persons lost to follow-up (very short survival for married persons, very long survival times for those not married) did not appear to influence our findings.

Additional cohort studies are needed that examine relationship, age, and gender of household members, as well as examine the dynamics of change in marital status and living arrangements in relation to mortality. It is also important to study the length of time in a specific living arrangement to address whether changes in marital status and living arrangement have acute effects leading to death or comprise a chronically stressful situation affecting health and mortality over time.

In conclusion, we found a stronger pattern of association between living arrangements and survival for men than for women, and for middle-aged men than for older men. There was no additional survival advantage for persons who lived with someone other than a spouse compared with those living alone, and—particularly for men—the critical factor for survival was the presence of a spouse. □

Acknowledgments

This work was supported by grant 5R37 AG05284 from the National Institute on Aging. Preliminary findings in this report were presented at the 118th meeting of the American Public Health Association in New York City on October 4, 1990. The authors wish to thank Dr. Walter Ettinger for gerontological advice, David Lein and John Barclay for assistance with data processing, and Christine Choy for assistance with preparation of the manuscript.

References

1. US Dept of Commerce, Bureau of the Census. Marital status and living arrangements: March 1985. *Current Population Reports, Population Characteristics,* Series P-20, No. 410. Washington, DC: US Govt Printing Office; 1986.
2. US Dept of Commerce, Bureau of the Census. Marital status and living arrangements: March 1989. *Current Population Reports, Population Characteristics,* Series P-20, No. 445. Washington, DC: US Govt Printing Office; 1990.
3. Siegel JS, Taeuber CM. Demographic perspectives on the long-lived society. *Daedalus.* 1986;115:77–117.
4. National Center for Health Statistics. *Mortality from Selected Causes by Marital Status: United States, Parts A & B.* Rockville, Md: National Center for Health Statistics; 1970. Public Health Service Publication 1000, Series 20, No. 8a, 8b.
5. Korbin FE, Hendershot GE. Do family ties reduce mortality? Evidence from the United States, 1966–1968. *J Marriage Fam.* 1977;39:737–745.
6. Trovato F, Lauris G. Marital status and mortality in Canada: 1951–1981. *J Marriage Fam.* 1989;51:907–922.
7. Hu Y, Goldman N. Mortality differentials by marital status: an international comparison. *Demography.* 1990;27:233–250.
8. Seeman TE, Kaplan GA, Knudsen L, Cohen R, Guralnik J. Social network ties and mortality among the elderly in the Alameda County Study. *Am J Epidemiol.* 1987;126:714–723.
9. Berkman LF, Syme SL. Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. *Am J Epidemiol.* 1979;109:186–204.
10. House JS, Robbins C, Metzner HL. The association of social relationships and activities with mortality: prospective evidence from the Tecumseh Community Health Study. *Am J Epidemiol.* 1982;116:123–140.
11. Schoenbach VJ, Kaplan BH, Fredman L, Kleinbaum DG. Social ties and mortality in Evans County, Georgia. *Am J Epidemiol.* 1986;123:577–591.
12. Gove WR. Sex, marital status, and mortality. *Am J Sociol.* 1973;79:45–67.
13. Helsing KJ, Szklo M. Comstock GW. Factors associated with mortality after widowhood. *Am J Public Health.* 1981;71:802–809.
14. Commonwealth Fund Commission. Aging

alone: profiles and projections. *The Commonwealth Fund.* 1988:1–83.
15. Wilson AJE. Sociological aspects of nutrition and aging. In: Hsu J, Davis R, eds. *Handbook of Geriatric Nutrition.* Park Ridge, NJ: Noyes Publications; 1981:43–55.
16. Select Committee on Nutrition and Human Needs of the US Senate. Nutrition and the aged. In: *Nutrition and Human Needs,* part 14. Washington, DC: US Govt Printing Office; 1969.
17. US Dept of Health, Education, and Welfare. *Older Americans Act of 1965, as Amended, and Related Acts.* Washington, DC: US Govt Printing Office; 1978.
18. Wingard DL, Cohn BA, Kaplan GA, Cirillo PM, Cohen RD. Sex differentials in morbidity and mortality risks examined by age and cause in the same cohort. *Am J Epidemiol.* 1989;130:601–610.
19. Kaplan GA, Haan MN. Is there a role for prevention among the elderly? Epidemiological evidence from the Alameda County Study. In: Ory MG, Bond K, eds. *Aging and Health Care: Social Science and Policy Perspectives.* London, England: Taristock Publications; 1989:27–51.
20. Feldman JJ, Makuc DM, Kleinman JC, Cornoni-Huntley J. National trends in educational differentials in mortality. *Am J Epidemiol.* 1989;129:19–33.
21. Berkman LF. Maintenance of health, prevention of disease, a psychosocial perspective. In: *Health of an Aging America: Issues on Data for Policy Analysis. Vital and Health Statistics,* Series 4, No. 25. Hyattsville, Md: US Dept of Health and Human Services; 1988. DHHS publication PHS 89–1488.
22. Kaplan GA, Seeman TE, Cohen RD, Knudsen LP, Guralnik J. Mortality among the elderly in the Alameda County Study: behavioral and demographic risk factors. *Am J Public Health.* 1987;77:307–312.
23. Kotler P, Wingard DL. The effect of occupational, marital and parental roles on mortality: the Alameda County Study. *Am J Public Health.* 1989;79:607–612.
24. Sorlie PD, Rogot E. Mortality by employment status in the national longitudinal mortality study. *Am J Epidemiol.* 1990;132:983–992.
25. National Institutes of Health Concensus Development Panel on the Health Implication of Obesity. Health implications of obesity. *Ann Intern Med.* 1985;103:1073–1077.
26. Camacho TC, Kaplan GA, Cohen RD. Alcohol consumption and mortality in Alameda County. *J Chron Dis.* 1987;40:229–236.
27. Rotevatn S, Akslen LA, Bjelke E. Lifestyle and mortality among Norwegian men. *Prev Med.* 1989;18:433–443.
28. Campbell AJ, Diep C, Reinken J, McCosh L. Factors predicting mortality in a total population sample of the elderly. *J Epidemiol Community Health.* 1985;39:337–342.
29. Seeman TE, Guralnik JM, Kaplan GA, Knudsen L, Cohen R. The health consequences of multiple morbidity in the elderly: the Alameda County Study. *J Aging Health.* 1989;1:50–66.
30. Kovar MG. Aging in the eighties, age 65 years and over and living alone, contacts with family, friends, and neighbors: preliminary data from the Supplement on Aging to the National Health Interview Survey: United States, January–June 1984. In: *Vital and Health Statistics,* No. 116, Hyattsville, Md: US Dept of Health and Human Services; 1986. DHHS publication 86–1250.
31. Magaziner J, Cadigan DA, Hebel JR, Parry RE. Health and living arrangements among older women: does living alone increase the risk of illness? *J Gerontol.* 1988;43:M127–M133.
32. Worobey JL, Angel RJ. Functional capacity and living arrangements of unmarried elderly persons. *J Gerontol.* 1990;45:S95–S101.
33. Davis MA, Randall E, Forthofer RN, Lee ES, Margen S. Living arrangements and dietary patterns of older adults in the United States. *J Gerontol.* 1985;40:434–442.
34. Ryan AS, Martinez GA, Wysong JL, Davis MA. Dietary patterns of older adults in the United States: NHANES II 1976–1980. *Hum Biol.* 1989;1:321–330.
35. Davis MA, Murphy SP, Neuhaus JM, Lein D. Living arrangements and dietary quality of older US adults. *J Am Diet Assoc.* 1990;90:1667–1672.
36. Miller HW. Plan and operation of the Health and Nutrition Examination Survey: United States 1971–1973. *Vital and Health Statistics,* Series 1, No. 10a. Hyattsville, Md: National Center for Health Statistics; 1978. US Dept of Health, Education, and Welfare publication PHS 79–1310.
37. National Center for Health Statistics. Plan and operations of the HANES I Augmentation Survey of adults 25–74 years: United States 1974–1975. *Vital and Health Statistics,* Series 1, no. 14. Hyattsville, Md: National Center for Health Statistics; 1978. US Dept of Health, Education, and Welfare publication PHS 78–1314.
38. National Center for Health Statistics. Plan and operation of the NHANES I Epidemiologic Follow-up Study 1982–84. *Vital and Health Statistics,* Series 1, No. 22. Hyattsville, Md: National Center for Health Statistics; 1987. US Dept of Health and Human Services publication PHS 87–1324.
39. Madans JH, Kleinman JC, Cox CS, et al. 10 years after NHANES I: report of initial follow-up, 1982–84. *Public Health Rep.* 1986;101(5):465–473.
40. Machlin SR, Kleinman JC, Madans JH. Validity of mortality analysis based on retrospective smoking information. *Stat Med.* 1989;8:997–1009.
41. Kleinman JC, Donahue RP, Harris MI, Finucane FF, Madans JH, Brock DB. Mortality among diabetics in a national sample. *Am J Epidemiol.* 1988;128:389–401.
42. National Center for Health Statistics. Health statistics on older persons: United States, 1986. *Vital and Health Statistics,* Series 3, no. 25. Hyattsville, Md: National Center for Health Statistics; 1987. US Dept of Health and Human Services publication PHS 87–1409.
43. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data.* New York, NY: John Wiley; 1980:70–118.
44. SAS Institute Inc. *SUGI Supplemental Library User's Guide,* version 5 ed. Cary, NC: SAS Institute Inc; 1986:437–466.
45. Davis MA, Lee ES, Neuhaus JM. Living arrangements, sense of well-being, and use of mental health services by US adults: implication for health policy. *J Health Hum Resources Administration.* 1987;9:319–339.
46. Hanson BS, Isacsson SO, Janzon L, Lindell SE. Social network and social support influence mortality in elderly men. *Am J Epidemiol.* 1989;130:100–111.
47. Berkman LF. Social networks, support, and health: taking the next step forward. *Am J Epidemiol.* 1986;123:559–562.
48. Blazer DG. Social support and mortality in an elderly community population. *Am J Epidemiol.* 1982;115:684–694.
49. Tayback M, Kumanyika S, Chee E. Body weight as a risk factor in the elderly. *Arch Intern Med.* 1990;150:1065–1072.
50. Gore S. Stress-buffering functions of social supports: an appraisal and clarification of research models. In: Dohrenwend BS, Dohrenwend BP, eds. *Stressful Life Events and Their Contexts.* New York, NY: Prodist; 1981:202–222.
51. Rosengren A, Wedel H, Wilhelmsen L. Marital status and mortality in middle-aged Swedish men. *Am J Epidemiol.* 1989;129:54–64.

Readmission after Surgery in Washington State Rural Hospitals

H. Gilbert Welch, MD, MPH, Eric H. Larson, MS, L. Gary Hart, PhD, and Roger A. Rosenblatt, MD, MPH

ABSTRACT

Background. Because of concern about the quality of care in rural hospitals, we examined readmission following four surgical procedures commonly performed in Washington State rural hospitals: appendectomy, cesarean section, cholecystectomy, and transurethral prostatectomy.

Methods. In a retrospective cohort study, we identified all patients discharged after receiving one of the foregoing procedures using the statewide hospital discharge database. Readmissions to any hospital in the state within 7 or 30 days of discharge were also identified.

Results. During the 2-year period examined, there were no significant differences in readmission rates for surgeries performed in rural and urban hospitals, although the readmission rates for all four procedures were nominally lower in rural hospitals. Logistic regression analyses that controlled for factors that influence readmission did not change these results.

Conclusions. Investigating readmission rates following common surgeries, we found no evidence of low-quality surgical care in Washington State rural hospitals. Early readmission is an imperfect marker for poor surgical outcome, however, and other proxies for quality remain to be examined. (*Am J Public Health.* 1992;82:407–411)

Introduction

Rural hospitals account for almost half of all short-stay nonfederal hospitals in the United States.[1] They facilitate access to basic medical services in geographically isolated areas. Equally important is their role of providing emergency care, particularly for cases of accidental injury,[2] to residents who would otherwise experience long delays in traveling to distant facilities. The finding that 10% of the nation's rural hospitals have closed during the past decade has therefore been the subject of considerable concern.[3]

One cause of rural hospital closures has been an increasing centralization of medical services into urban hospitals. This phenomenon is accelerated by assertions that small rural facilities deliver care of inferior quality.[4,5] Rural residents may be acting on this belief when deciding to seek elective care in large urban hospitals.[6,7] Because diminished use undermines the fiscal viability of any hospital, assertions of low-quality care in rural hospitals may threaten their ability to survive. Unfortunately, this area has been inadequately studied.[2]

To investigate one aspect of quality, we examined patients who underwent any of four surgical procedures frequently performed in Washington State rural hospitals. We tested the hypothesis that rural hospitals had the same readmission rates, following these surgeries, as their urban counterparts. Readmission rates are an often used, although imperfect, proxy for quality.[8–10] Readmissions to any Washington State hospital within 7 or 30 days of being discharged were analyzed.

Methods

Readmission Data

Data were obtained from the Washington State Commission Hospital Abstract Reporting System (CHARS). The CHARS database used in this analysis includes information on all discharges from Washington State hospitals (excluding newborns) occurring during calendar years 1987 and 1988. The methodology used to link the 925 309 discharge records into 636 328 individual patient records has been described elsewhere.[11] Patients transferred from one hospital to another were not counted as readmissions.

Definition of Rural

Although there are many schemes that differentiate rural and urban areas,[12] all are imperfect attempts to dichotomize a multidimensional attribute. Recognizing this problem, we chose the current criteria of the Washington State Department of Health for defining rural hospitals (V. Gibbs, personal communication, December 1989). Rural hospitals have been de-

H. Gilbert Welch is with the Departments of Medicine and Community and Family Medicine, Department of Veterans Affairs Medical Center, White River Junction, Vermont, and Dartmouth Medical School, Hanover, New Hampshire. Eric H. Larson, L. Gary Hart, and Roger A. Rosenblatt are with the Department of Family Medicine, Washington, Alaska, Montana, and Idaho (WAMI) Rural Health Research Center, University of Washington, Seattle.

Requests for reprints should be sent to H. Gilbert Welch, MD, MPH, Medical Service, White River Junction Veterans Administration Hospital, North Hartland Road, White River Junction, VT 05009.

This paper was submitted to the journal December 11, 1990, and accepted with revisions June 10, 1991.

TABLE 1—Discharge Diagnosis-Related Groups Selected for Analysis

Procedure Description (ICD-9 procedure code)	Diagnosis-Related Group
Appendectomy (47.0)	
With complicated principal diagnosis with CC	164
With complicated principal diagnosis without CC	165
Without complicated principal diagnosis with CC	166
Without complicated principal diagnosis without CC	167
Cesarean section (74.1)	
With CC	370
Without CC	371
Cholecystectomy (51.22)	
With common duct exploration with CC	195
With common duct exploration without CC	196
Without common duct exploration with CC	197
Without common duct exploration without CC	198
Transurethral prostatectomy (60.2)	
With CC	336
Without CC	337

Note. Comorbid conditions must exist prior to the admission, while complications must develop during the hospitalization. CC = complications and/or comorbidity; ICD-9 = *International Classification of Diseases*, 9th version.

fined as those that are located in counties with a population of less than 200 000 and that are more than 5 miles from the boundaries of a city or city cluster with a population of more than 25 000. Designated hospitals are generally much farther away from cities, but the criteria were fashioned to allow specific hospitals near small cities in rural landscapes to qualify. Forty-four hospitals in Washington State are designated as rural using these criteria, while the other 62 hospitals are considered urban.

Selection of Common Surgical Procedures

Using the primary *International Classification of Diseases*, 9th version (ICD-9), procedure code, we identified the 10 most common surgical procedures performed in Washington's rural hospitals. Five of these procedures involved routine peripartum care: artificial rupture of membranes, episiotomy, spontaneous vaginal delivery, repair of obstetrical laceration, and circumcision. The other 5 involved more complex procedures and were selected for analysis because of their higher a priori potential for complications. One procedure, hysterectomy, was excluded because the two distinct surgical approaches (abdominal and vaginal) were found to be differentially performed in rural and urban areas. Two of the four procedures analyzed here are predominantly elective: cholecystectomy and transurethral prostatectomy. Two are more likely to be nonelective: appendectomy and cesarean section.

Selection of Study Population

All patients undergoing operations in Washington State hospitals after January 31, 1987, were eligible for inclusion in the analysis. Patients with previous admissions were excluded from study. Additionally, study patients had to be discharged (after surgery) on or before November 15, 1988, to be eligible. This end date allowed for all subsequent admissions within 30 days to be captured, provided they were less than 2 weeks in duration.

The analysis was restricted to patients with one of four ICD-9 primary operative procedure codes: appendectomy (47.0), cesarean section (74.1), cholecystectomy (51.22), or transurethral prostatectomy (60.2).[13] The analysis was further restricted to patients whose discharge was assigned one of the diagnosis-related groups (DRGs) listed in Table 1. This eliminated patients for whom the procedure was a relatively insignificant portion of hospitalization (e.g., a patient undergoing induction chemotherapy for leukemia who requires an appendectomy). This restriction eliminated 2.6% of the appendectomy patients, 0.2% of the cesarean section patients, 3.4% of the cholecystectomy patients, and 9.4% of the prostatectomy patients. No detectable rural–urban differences in in-hospital mortality were found for patients in these DRGs. The 60 patients who died within the hospitalization during which they were operated on (6 at rural hospitals, 54 at urban hospitals) were excluded from further analysis.

Analysis of Readmissions

Readmission was defined as a second admission, to any hospital in Washington State, within 7 or 30 days of discharge. Crude 7-day and 30-day readmission rates were compared in rural and urban hospitals. In addition, the risk of readmission for patients operated on in rural hospitals relative to those operated on in urban hospitals was calculated. Confidence intervals surrounding the relative risk estimates were calculated using the standard error approximation.[14]

The initial findings for each surgery were verified through analysis of more homogeneous subgroups (e.g., analysis of appendectomies was restricted to patients between 10 and 40 years of age, analysis of cesarean sections was restricted to patients between 20 and 30 years of age, and analyses of all procedures were restricted to patients with private insurance). Urban and rural differences were tested via the standard χ^2 corrected for continuity test and t test. In addition, multiple logistic regression was used to simultaneously control for factors other than the rural or urban location of surgery that are thought to influence readmission rates. The factors controlled for from the discharge abstracts included patient age and sex, payor type (e.g., Medicaid), emergency admission, length of stay, and presence of comorbid conditions.

This study is basically descriptive and includes the population of relevant surgeries and readmissions for a 2-year period in Washington State. Expanding the study temporally or geographically would increase the power of the analyses but would obscure precisely which population was being examined. The overall power of the study when the four surgeries are aggregated is .75 for a difference of 3 readmissions per 1000 admissions, but is lower for each of the surgeries taken separately. However, it should be remembered that study results are not extrapolated beyond their descriptive limits and that all relevant cases have been included in the analyses. Thus, the study's results stand on their own as a description of Washington State for 2 years but must be viewed cautiously if extrapolated further.

Results

During the study period, approximately 12% of all hospital discharges in Washington came from the state's 44 rural hospitals. These 44 hospitals were responsible for 10% of all procedures performed.

TABLE 2—Diagnoses Accounting for More than 5% of 7-day Readmissions for Four Surgical Procedures

Procedure	DRG	DRG Description	Readmissions, %	Cumulative %
Appendectomy				
	418	Postoperative infection	28	28
	415	OR procedure for infection	26	54
	183, 184	Gastroenteritis/digestive disorders	15	69
	453	Complications of treatment	5	74
Cesarean section				
	376	Postpartum complications	68	68
	377	Postpartum complications, to OR	8	76
Cholecystectomy				
	418	Postoperative infection	12	12
	182, 183	Gastroenteritis/digestive disorders	12	24
	452, 453	Complications of treatment	11	35
	188, 189	Other digestive system diagnoses	11	46
	204	Pancreatic disorders	6	52
	442	Operation for hemorrhage/repair	6	58
Transurethral Prostatectomy				
	452, 453	Complications of treatment	36	36
	331	Other kidney/urinary tract diagnoses	6	42

Note. DRG = diagnosis-related group; OR = operating room. Paired DRGs (188, 189 and 452, 453) combine patients with and without complications and/or comorbidities (CC). DRGs 182 (without CC) and 183 (with CC) are for patients over age 17; DRG 184 is for patients 1–17 years of age.

TABLE 3—Rural and Urban Readmissions following Surgery

	Appendectomy		Cesarean Section		Cholecystectomy		Transurethral Prostatectomy	
	Rural Hospitals	Urban Hospitals	Rural Hospitals	Urban Hospitals	Rural Hospitals	Urban Hospitals	Rural Hospitals	Urban Hospitals
No. of cases	1280	6022	2452	18 013	1639	8378	1116	6862
Age–sex adjusted rate (per 1000 population)	.81	.78	210[a]	216[a]	1.27	1.03	1.71	1.53
Mean age	25.9	26.7	26.2**	27.5	51.8	51.1	70.8	70.4
Male, %	60.4	58.3	...	...	25.4	25.0	...	...
Medicaid/self-pay, %	36.1**	22.5	45.4**	27.5	21.8**	13.6	2.3	2.5
Medicare, %	5.7**	3.9	0.3	0.1	29.8**	24.7	77.8**	68.6
Mean length of stay	3.5*	3.7	3.7**	4.3	5.3	5.2	3.2	3.3
Length of stay outliers, %	0.7	0.9	0.4	0.7	0.8	0.8	1.0	0.8
7-day readmission rate (per 1000 procedures)	8.59	12.45	4.07	5.94	7.37	13.13	12.54	17.48
30-day readmission rate (per 1000 procedures)	19.53	27.64	9.78	12.21	30.50	32.46	37.63	44.30
7-day relative risk (95% confidence interval)	.73	(0.42, 1.27)	.71	(0.39, 1.29)	.60	(0.35, 1.03)	.74	(0.45, 1.22)
30-day relative risk (95% confidence interval)	.76	(0.52, 1.10)	.82	(0.56, 1.20)	.95	(0.73, 1.23)	.86	(0.65, 1.15)

[a]For cesarean section, age adjusted rates are expressed as procedures per 1000 deliveries.
*$P < .05$, intrasurgery rural/urban comparison.
**$P < .01$, intrasurgery rural/urban comparison.

Of the four surgical procedures analyzed here, 14% were performed in rural hospitals.

The most common 7-day readmission diagnoses following each surgery are listed in Table 2. The dominance of diagnoses such as postoperative infection, digestive disorders, and complications of treatment in readmitted patients indicates that readmission within the 7-day period generally reflects complications of surgery. Careful examination of less frequent readmission diagnoses revealed additional conditions compatible with complications from recent surgery.

Procedure-specific data are shown in Table 3. With the exception of cesarean section, age–sex-adjusted procedure rates are higher for the population living in the service areas of rural hospitals. Patients hospitalized in rural hospitals were significantly more likely to be uninsured or receive Medicaid (except in the case of transurethral prostatectomy, where the majority of patients are eligible for Medicare). For two procedures (appendectomy and cesarean section), the mean length of stay for patients hospitalized in rural hospitals was significantly shorter than for those hospitalized in urban centers; for the remaining two procedures, no difference in length of stay was observed.

Figure 1 displays the 7-day readmission rates for each procedure in both rural and urban hospitals. Regardless of the procedure examined, readmission rates follow-

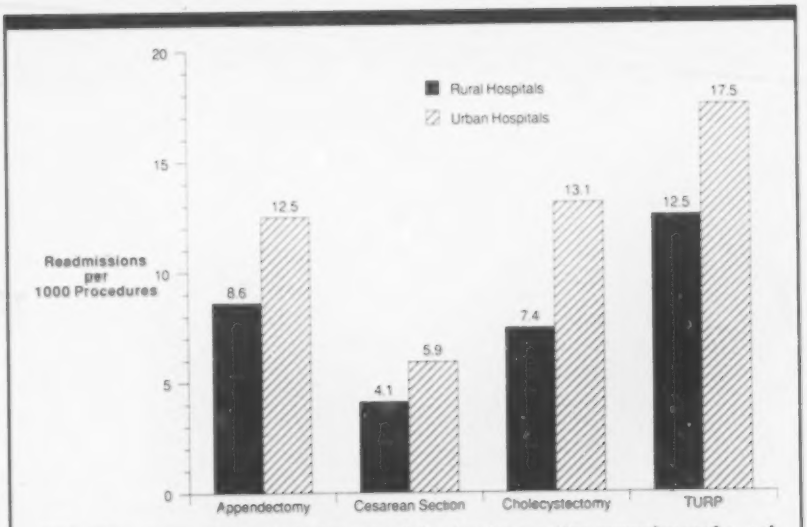


FIGURE 1—Seven-day readmission rates following four routine surgeries performed in Washington State rural and urban hospitals (1987–1988). TURP = transurethral prostatectomy.

ing surgeries performed in rural hospitals were lower than for those performed in urban hospitals. The 7-day relative risk of readmission for patients operated on in rural hospitals (shown in Table 3) ranged from .60 for cholecystectomy to .74 for transurethral prostatectomy. Thirty-day relative risk estimates were somewhat higher and ranged from .76 for appendectomy to .95 for cholecystectomy. Relative risk estimates were consistently below one in analyses of more homogeneous subgroups.

To further examine the potential confounding influence of case mix, a logistic regression analysis was performed. This analysis was designed to test, for each of the four surgeries, whether the finding of lower readmission rates following surgery in rural hospitals would persist after simultaneously controlling for other factors that influence rehospitalization. The variables included in the regressions were patient age, sex, payor type (e.g., Medicaid/self-pay, Medicare, and other insurance), emergency admission, length of stay, and number of comorbid conditions. Alternative functional forms of several of these factors were tested for both 7-day and 30-day readmissions. Many of the control variables were significantly associated with readmission. Although not statistically significant, the regression coefficient for surgery in rural hospitals remained negative for all procedures.

Discussion

We investigated four common surgeries in Washington State hospitals and found that rural hospitals did not have elevated readmission rates. In fact, for each of the four procedures, 7-day readmission rates for rural hospitals were consistently lower than those in urban hospitals, although not significantly so. When all readmission data are pooled, however, the statistical power of this investigation to detect significantly different rural readmission rates (were they to exist) increases our confidence in its major finding. Furthermore, this descriptive study includes every event for a 2-year period in the population of interest. Expanding the study in time or to other locales might make the conclusion more certain but would obscure precisely which population was being studied.

Readmission rates following surgery are an indirect measure of quality. The diagnoses commonly observed during readmission (postoperative infection, digestive disorders, and complications of treatment), however, suggest that readmission generally reflects complications of surgery. Some patients develop complications immediately following surgery and remain hospitalized longer, rather than being discharged and readmitted. This may reflect important quality problems that cannot be identified using readmission as a marker. The use of logistic regression analysis, however, controlled for this potentially confounding variable.

The possibility that case-mix differences explain outcome differences is always a concern in investigations of this type. There are two major reasons one might posit that Washington's urban hospitals had "sicker patients." One possibility is that sicker patients living in rural areas tend to have their surgery done in urban hospitals. Yet, removing rural residents whose surgery was performed in urban hospitals had a minimal effect on the analysis. A second possibility is that urban hospitals serve a more disadvantaged population than do rural hospitals. Inner-city patients may be discharged to a more hostile home environment, making readmission more likely. However, eliminating all Medicaid, self-pay, and charity patients (the CHARS data contain no information on patient race, income, or education) from the patient population studied or controlling for factors such as age, payor type, and comorbid conditions also had minimal effect. To the extent case mix can be imputed from the discharge database, these analyses suggest that it is not an important confounder in this study.

Finally, it is possible that our findings reflect differences in patient or provider behavior, obscuring real quality differences. Rural patients, for example, may be more hesitant than urban patients to seek care for postsurgical problems. It is not possible to ascertain the effect that patient attitudes might have on the results. Alternatively, rural providers may be more likely to treat postsurgical problems on an outpatient basis and have a higher threshold for readmission than urban providers. However, the extended distances that patients often travel to rural hospitals suggest that rural providers would be more likely to display the opposite behavior.

For these four common surgical procedures, it is hard to argue that Washington State's rural hospitals performed poorly in comparison to their urban counterparts. In nominal terms, rural readmission rates were, in fact, lower. The extent to which our findings extend to other commonly provided surgical services is unknown. Although these data are from a limited geographic region and may not reflect patterns elsewhere, similar work has demonstrated that high-quality obstetrical care can be provided in rural settings.[15,16] Our study suggests that these four routine surgical procedures can be provided at rural hospitals with a quality comparable to urban facilities. This kind of information can help administrators and physicians in rural hospitals define a range of services that enhances both fiscal performance and operator proficiency without compromising quality. More important, it serves to discredit the general perception that rural hospitals provide inferior care. □

Acknowledgments

Dr. Welch was a Veterans Administration (VA) Robert Wood Johnson clinical scholar at the Seattle VA Medical Center when this study was initiated and is now supported by a research associate award from the VA Health Services Research and Development Career Development Program. The investigation was supported by a grant to the WAMI Rural Health Research Center of the University of Washington School of Medicine from the Office of Rural Health Policy, Health Resources and Services Administration (HAR000017-02), Public Health Service.

References

1. Hart LG, Rosenblatt RA, Amundson BA. Is there a role for the small rural hospital? *J Rural Health.* 1990;6:101–118.
2. Erman DA. Rural health care: the future of the hospital. *Med Care Rev.* 1990;47:33–73.
3. Merrlis M. *Rural Hospitals.* Washington, DC: Congressional Research Service; 1989.
4. Rohrer JE. Closing rural hospitals: reducing "institutional bias" or denial of access? *J Public Health Policy.* 1989;10:353–358.
5. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med.* 1979;301:1364–1369.
6. Mullner RM, McNeil D. Rural and urban hospital closures. *Health Aff.* 1986;5:131–141.
7. Amundson BA, Hughes RD. Are dollars really the issue for the survival of rural health services? *Rural Health Working Paper Series.* 1989;1(3).
8. Riley G, Lubitz J. Outcomes of surgery in the Medicare aged population: rehospitalization after surgery. *Health Care Financing Rev.* 1986;8:23–34.
9. Roos LL, Sharp SM. Becoming more efficient at outcomes research. *Int J Technol Assess Health Care.* 1988;4:555–571.
10. Holloway JJ, Medendorp SV, Bromberg J. Risk factors for early readmission among veterans. *Health Services Res.* 1990;25:213–237.
11. Welch HG, Larson EH. Patients requiring at least five admissions in one year—data from Washington State. *Med Care.* 1991;29:578–82.
12. Hewitt M. *Defining "Rural" Areas: Impact on Health Care Policy and Research.* Washington, DC: Office of Technology Assessment; 1989. Staff Paper.
13. US Department of Health and Human Services. *International Classification of Diseases—9th Version: Volume 3—Procedures.* Washington, DC: US Department of Health and Human Services; 1980. DHHS publication no. PHS 80-1260.
14. Matthews DE, Farewell VT. Using and understanding medical statistics. Basel, Switzerland: Karger; 1988:145–146.
15. Rosenblatt RA, Mayfield JA, Hart LG, Baldwin LM. Outcomes of regionalized perinatal care in Washington state. *West J Med.* 1988;149:98–102.
16. Rosenblatt RA, Reinken J, Shoemack P. Is obstetrics safe in small hospitals? Evidence from New Zealand's regionalized perinatal system. *Lancet.* 1985;24:429–431.

New Pathways in Epidemiology to Be Explored in Sydney

The International Epidemiological Association announces the following meeting:

New Pathways in Epidemiology
Thirteenth Scientific Meeting
University of Sydney, Australia
September 26 to 30, 1993

This exciting opportunity to participate in a major international scientific meeting will provide an opportunity to examine new pathways in global health, clinical practice, environmental issues, health services management, and methodology and also to savor Sydney in the Southern Hemisphere spring.

If you would like to be included on the conference mailing list, please send contact details—your name, title, affiliation/institution, and the best address for reaching you—to IEA Conference Secretariat, PO Box 746, Turramurra 2074 NSW, Australia.

The Behavioral Risk Factor Survey and the Stanford Five-City Project Survey: A Comparison of Cardiovascular Risk Behavior Estimates

Christine Jackson, PhD, Darius E. Jatulis, MS, and Stephen P. Fortmann, MD

ABSTRACT

Background. Nearly all state health departments collect Behavioral Risk Factor Survey (BRFS) data, and many report using these data in public health planning. Although the BRFS is widely used, little is known about its measurement properties. This study compares the cardiovascular risk behavior estimates of the BRFS with estimates derived from the physiological and interview data of the Stanford Five-City Project Survey (FCPS).

Method. The BRFS is a random telephone sample of 1588 adults aged 25 to 64; the FCPS is a random household sample of 1512 adults aged 25 to 64. Both samples were drawn from the same four California communities.

Results. The surveys produced comparable estimates for measures of current smoking, number of cigarettes smoked per day, rate of ever being told one has high blood pressure, rate of prescription of blood pressure medications, compliance in taking medications, and mean total cholesterol. Significant differences were found for mean body mass index, rates of obesity, and, in particular, rate of controlled hypertension.

Conclusions. These differences indicate that, for some risk variables, the BRFS has limited utility in assessing public health needs and setting public health objectives. A formal validation study is needed to test all the risk behavior estimates measured by this widely used instrument. (*Am J Public Health.* 1992;82:412–416)

Introduction

Nearly all state health departments now collect Behavioral Risk Factor Survey (BRFS) data,[1] and many report using BRFS data to set state objectives, plan statewide health education programs, and support public health legislation.[2,3] But although the BRFS is widely used, little is known about its measurement properties. The Centers for Disease Control has published annual reports of BRFS prevalence estimates for selected states[4,5] and national prevalence estimates from pooled data for cigarette smoking,[6] overweight,[7] and alcohol use.[8] A study of the test-retest reliability of the BRFS reported estimates that were generally 0.60 or greater for cardiovascular risk behaviors other than dietary behaviors, for which lower estimates were found.[9] Studies comparing BRFS estimates to other measures are limited to one study that found that, for selected risk behaviors, total population estimates from the combined 1981–1983 BRFS data were similar to those from national surveys conducted between 1979 and 1980,[10] and to a second study that found that most age- and sex-specific estimates for smoking and binge drinking were comparable for the BRFS and an in-person interview.[11]

The present study builds on this previous work by comparing the BRFS cardiovascular risk behavior prevalence estimates with estimates from the Stanford Five-City Project's random household survey (FCPS), which is the principal evaluation instrument of the Stanford Five-City Project.[12] Importantly, the FCPS includes physiological comparison measures for cigarette smoking, blood pressure, blood cholesterol, height, and weight. These measures currently provide the best estimates of risk and so can be used as benchmarks against which the BRFS estimates can be compared. This study also compares BRFS variables not previously evaluated, including smoking rate, body mass index (BMI), and total cholesterol.

Method

Study Samples

Both the BRFS and FCPS samples were drawn from the same four northern California communities participating as education or control cities in the Stanford Five-City Project[12] (see Table 1).

The BRFS is the first of three biennial random telephone surveys planned for the Five-City Project. The Waksberg[13] random-digit dialing method was used to generate the probability sample, and a detailed description of the sampling method and contact protocol for this study is available elsewhere.[14] The total sample included 2360 individuals, of whom 1588 (67%) were aged 25 to 64, spoke English, and were included in the present study. A response rate of 83% was obtained for the total sample. Because confirmation of the telephone number and city of residence were the only data obtained from nonre-

The authors are with the Stanford Center for Research in Disease Prevention in Palo Alto, CA.

Requests for reprints should be sent to Dr. Christine Jackson, Stanford Center for Research in Disease Prevention, 1000 Welch Road, #104, Stanford University School of Medicine, Palo Alto, CA 94304-1885.

This paper was submitted to the journal August 31, 1990, and accepted with revisions June 10, 1991.

spondents, it is not possible to calculate a response rate for the 25- to 64-year-old subsample.

The FCPS is the fourth biennial random cross-sectional survey conducted by the Five-City Project, and a detailed description of the survey method is available.[12] The total sample included 2378 individuals, of whom 1512 (64%) were aged 25 to 64, spoke English, and were included in the present study. The response rate was 56% for the total sample and 61% for the 25- to 64-year-old subsample.

Measures

Components of the BRFS and FCPS were selected for comparison if an interview measure from the BRFS and a physiological measure from the FCPS purport to measure the same variable (e.g., current smoking status, weight), or if interview questions from the BRFS and FCPS have the same or similar wording and so appear to measure the same variable (e.g., number of cigarettes smoked per day, compliance in taking blood pressure medications). An abbreviated description of the comparison is presented in Table 2; a detailed description of the physiological assessments used by the Five-City Project is available,[12] and a copy of the wording of the BRFS and FCPS interview measures compared in this study may be requested from the first author.

Analyses

Continuous measures for the two surveys were compared using *t* tests; when the sample variances were found to be significantly different, a separate variance *t* test was used. The chi-square test was used to evaluate significance for most dichotomous outcomes.

To control for possible effects of demographic differences between samples, adjusted total sample estimates were also computed for each outcome. Adjusted values for the continuous outcomes were found using analysis of covariance; logistic regression was used for categorical outcomes.[20]

Results

Sample Characteristics

Compared with the FCPS sample, the BRFS sample had more females (a result consistent with the generally higher rate of refusal among males for telephone surveys)[21] and a lower mean age, but the same mean educational level (see Table 1).

TABLE 1—Survey and Sample Characteristics of the Behavioral Risk Factor Survey (BRFS) and the Five-City Project Survey (FCPS)

	BRFS (1987)	FCPS (1986)	
Geographic area	education and control cities of the Five-City Project	education and control cities of the Five-City Project	
Sample method	random-digit dialing	random household	
Survey method	telephone interview	face-to-face interview; physiological assessment	
Eligibility	one person aged 18 or older	all persons aged 12–74 in household	
Response rate	83% (total sample) NA (ages 25–64)	56% (total sample) 61% (ages 25–64)	
Sample size	2360 (total sample) 1588 (ages 25–64)	2378 (total sample) 1512 (ages 25–64)	
Sample demographics			*P*
% Male	41.6 (n = 660)	46.2 (n = 698)	
% Female	58.3 (n = 926)[a]	53.8 (n = 814)	.01
% Age 25–34	40.2 (n = 638)	34.2 (n = 517)	
% Age 35–44	28.1 (n = 446)	28.2 (n = 427)	
% Age 45–54	15.6 (n = 248)	18.8 (n = 284)	
% Age 55–64	16.1 (n = 256)	18.8 (n = 284)	.002
Mean age	40.1 (SD = 11.4)	41.5 (SD = 11.5)	.001
% <12 years education	9.4 (n = 150)	8.9 (n = 135)	
% 12 years education	23.0 (n = 365)	25.9 (n = 391)	
% 13–15 years education	35.3 (n = 560)	29.7 (n = 449)	
≥16 years education	32.3 (n = 513)	35.3 (n = 534)	.01
Mean years education	14.2 (SD = 2.2)	14.1 (SD = 2.9)	.12

Note. SD = standard deviation of a sample.
[a]Gender data missing for two respondents.

Cigarette Smoking Estimates

The BRFS and FCPS provide comparable estimates of current smoking rates and mean number of cigarettes smoked per day, with none of the between-survey differences being statistically significant (see Table 3). The results are less consistent for rates of heavy smoking. Although the FCPS estimate is significantly higher for the total sample comparison, most of this effect is attributable to the City 3 subsample, with the remaining cities having comparable estimates of heavy smoking. As this measure is identically worded on both surveys and there are no known events or method differences unique to City 3, it is not clear why this discrepancy occurred.

Blood Pressure Estimates

Generally comparable estimates were found for the measures of ever being told one has high blood pressure and of being prescribed blood pressure medications (Table 4). Although significant between-survey differences were found for some cities, neither survey is systematically higher or lower and the differences between the total sample estimates are not significant.

The BRFS and FCPS also measure compliance in taking blood pressure medications (data not shown). Both surveys found high levels of compliance, with 95% and 93% of the BRFS and FCPS samples, respectively, reporting that they always or almost always take their blood pressure medications. While these estimates indicate that the surveys provide comparable estimates of compliance, they also indicate that both surveys probably overestimate actual compliance rates. In a recent review of medical compliance studies, for example, Kaplan and Simons[22] report that most studies find at least 33% noncompliance for therapeutic regimens.

Of all the outcomes compared in this study, the between-survey differences are largest for estimates of controlled hypertension among individuals with hypertensive medications prescribed. The magnitude of the between-survey differences, ranging from 16% to 62%, suggests that a true measurement difference exists and raises some concern about the utility of this BRFS item. The 1987 California BRFS* found that 91% of hypertensives aged 25 to 64 report having their high

*Unpublished 1987 BRFS tabulations from the Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion.

TABLE 2—Comparison Measures from the Behavioral Risk Factor Survey (BRFS) and the Five-City Project Survey (FCPS) for Comparable Variables

Variable	BRFS	FCPS
Cigarette smoking	(1) self-reported smoking status	self-reported smoking status verified by thiocyanate and carbon monoxide test[15]
	(2) self-reported number of cigarettes smoked per day	same
	(3) percent of heavy smokers, defined as those reporting 25 or more cigarettes smoked per day	same
Hypertension	(4) self-reported hypertension measured as ever being told by a doctor or other health professional that one has high blood pressure	same
	(5) self-reported medical prescription for hypertension control	self-reported medical prescription for hypertension control, verified by interviewer inspection of medications and recording of drug names
	(6) self-reported compliance in taking hypertension medications	same
	(7) self-reported blood pressure status as normal or under control	normotensive status, defined as systolic pressure < 140 mm and diastolic pressure < 90 mm, using automated blood pressure readings[16,17]
Overweight	(8) mean BMI, using self-reported weight and height	mean BMI, using weight measured with a balance scale and height measured with a metal rule
	(9) same BMI values as the FCPS, but based on self-reported height and weight	percent with obese BMI values, defined as being more than 130% of the mean BMI for males (BMI > 28.08) and females (BMI > 27.56)[18]
Cholesterol	(10) mean self-reported total cholesterol value	mean plasma cholesterol measured with the standardized Lipid Research Clinics protocol[19]

Note. BMI = body mass index.

blood pressure under control, an indication that the 89% estimate reported by the present sample is reliable. Considering the factors that could affect responses to the BRFS measure, it is most likely that the self-report format results in inflated estimates of control. It is possible, for example, that BRFS respondents who had lowered *either* their systolic or their diastolic pressure reported this as gaining control. In contrast, FCPS respondents must have had *both* systolic and diastolic pressures below physiological cut-points to be counted as having their high blood pressure under control. Overall, the data suggest that the BRFS self-report item actually measures the perception that one's blood pressure is under control and that the prevalence of this perception greatly exceeds real levels of controlled hypertension.

BMI Estimates

The BRFS estimates for both mean BMI and percent obese BMI were consistently and significantly lower than the FCPS estimates (see Table 5). The differences in mean BMI ranged from –1.0 to –2.2, with the –1.5 difference between the total sample estimates representing approximately a 10-lb difference in weight. Such underreporting of weight is to be expected, but the magnitude is larger here than has been observed elsewhere.[23,24]

Cholesterol Estimates

Before comparing the survey estimates of mean total cholesterol, we note that only 9% (n = 149) of the total BRFS sample reported ever being told a cholesterol value *and* were able to report a cholesterol value. Although this BRFS sample is relatively small, the mean cholesterol values are well within the normative range, and a demographic breakdown of the sample (not shown) reveals that expected increases in knowledge of one's cholesterol come as age and level of education increase, and that increases in mean total cholesterol come with age. Thus, despite its small size, this BRFS sample warrants cautious comparison with the FCPS sample. As shown in Table 5, the surveys produced comparable estimates of mean total cholesterol.

Adjusted Estimates

Adjusting the total sample estimates for age, sex, and education had no effect on the significance of the between-survey differences and had little effect on the actual prevalence estimates. (Because of the large difference in sample size, adjusted estimates were not computed for total cholesterol.) Significant survey × age or survey × education interactions were found for some outcomes. Examination of the age- and education-stratified estimates for all outcomes (not shown) indicate that the significant interactions were generally due to one age or educational subgroup being different from the rest, rather than to a complete reversal in the pattern of results.

Discussion

Comparing the cardiovascular risk behavior prevalence estimates for the total BRFS and FCPS samples, we found substantial between-survey agreement for several variables: current smoking rate, number of cigarettes smoked per day, rate of ever being told one has high blood pressure, rate of prescription of blood pressure medications, compliance in taking medications, and mean total cholesterol. These results indicate that, at least in large samples, the BRFS provides accurate estimates for these cardiovascular risk variables. The results also indicate that two of these variables, while producing comparable estimates, have other limitations. First, the estimates of compliance in taking blood pressure medications were found to be quite high and probably overestimate actual compliance rates. Second, only a small proportion of the BRFS sample could report a cholesterol value, indicating that, while this measure is useful in estimating knowledge of cholesterol, it has limited utility in estimating mean total cholesterol for a population.

TABLE 3—Behavioral Risk Factor Survey (BRFS) and Five-City Project Survey (FCPS): Estimates for Smoking

	% Current Smokers			Cigarettes per Day			% Heavy Smokers		
	BRFS (1588)	FCPS (1509)	Diff (SE)	BRFS (382)	FCPS (346)	Diff (SE)	BRFS (381)	FCPS (332)	Diff (SE)
All	24.8	23.0	1.8 (1.5)	18.7	18.4	0.3 (0.90)	23.4	30.1*	–6.8 (3.3)
City 1	23.8	23.7	0.1 (2.9)	17.9	17.0	0.9 (1.7)	20.4	26.5	–6.1 (6.1)
City 2	28.9	24.9	4.0 (3.3)	17.9	16.7	1.2 (1.8)	23.7	22.9	0.8 (6.1)
City 3	29.7	26.8	3.0 (3.2)	20.7	22.5	–1.8 (1.6)	26.9	42.9*	–16.0 (6.4)
City 4	15.4	15.4	0.1 (2.8)	17.4	15.5	1.9 (2.2)	20.0	21.7	–1.7 (8.1)
All adjusted	24.8[a]	23.0		18.7[b]	18.8		22.6[b]	30.1*	

Note. SE = standard error.
[a]adjusted for sex.
[b]adjusted for age, sex, and education.
*$P \leq .05$.

TABLE 4—Behavioral Risk Factor Survey (BRFS) and Five-City Project Survey (FCPS): Estimates for Blood Pressure

	% Ever Told of High BP			% Medication Prescribed			% BP Under Control		
	BRFS (1580)	FCPS (1507)	Diff (SE)	BRFS (338)	FCPS (346)	Diff (SE)	BRFS (122)	FCPS (135)	Diff (SE)
All	21.3	23.0	–1.7 (1.5)	36.7	39.0	–2.3 (3.7)	89.3	60.7***	28.6 (5.5)
City 1	18.9	20.9	–2.0 (2.7)	25.3	39.4*	–14.0 (7.2)	94.7	32.4***	62.3 (14.1)
City 2	23.5	27.2	–3.7 (3.2)	37.5	43.0	–5.5 (7.1)	94.3	77.5*	16.8 (8.2)
City 3	20.6	29.1**	–8.5 (3.0)	41.7	34.5	7.2 (6.9)	85.7	70.0	15.7 (9.7)
City 4	22.2	13.6**	8.6 (3.0)	41.8	41.9	–0.1 (9.3)	84.8	61.1	23.7 (12.4)
All adjusted	20.1[a]	20.4		34.9[b]	33.8		90.0[b]	60.2***	

Note. SE = standard error.
[a]adjusted for age and sex.
[b]adjusted for age, sex, and education.
*$P \leq .05$, **$P \leq .01$, ***$P \leq .001$.

TABLE 5—Behavioral Risk Factor Survey (BRFS) and Five-City Project Survey (FCPS): Estimates for Body Mass Index (BMI) and Cholesterol

	BMI			% Obese BMI			Cholesterol mmol/L		
	BRFS (1559)	FCPS (1501)	Diff (SE)	BRFS (1559)	FCPS (1501)	Diff (SE)	BRFS (149)	FCPS (1470)	Diff (SE)
All	24.2	25.7***	–1.5 (.16)	14.1	25.3***	–11.2 (1.4)	5.18	5.07	.11 (.12)
City 1	23.8	25.4***	–1.6 (.31)	11.6	21.8***	–10.2 (2.6)	4.97	5.08	–.11 (.16)
City 2	24.8	25.8**	–1.0 (.32)	17.8	25.2*	–7.4 (3.0)	4.92	5.15	–.23 (.19)
City 3	24.6	26.4***	–1.8 (.33)	15.5	31.6***	–16.0 (3.0)	5.53	5.11	.42 (.32)
City 4	23.7	25.1***	–1.4 (.34)	11.2	22.6***	–11.4 (2.9)	5.28	4.91	.37 (.26)
All adjusted	24.3[a]	25.8***		13.7[b]	24.7***		—	—	

Note. SE = standard error.
[a]adjusted for sex.
[b]adjusted for sex and education.
*$P \leq .05$, **$P \leq .01$, ***$P \leq .001$.

Comparable between-survey estimates were not found for the other risk behavior variables: prevalence of controlled blood pressure, mean BMI, and prevalence of obesity. For these variables, the total sample comparisons and most of the city-specific comparisons were statistically significant and of a magnitude likely to affect public health planning decisions. Of particular concern for those using BRFS data for needs assessment and program planning are the overestimation of controlled hypertension and the underestimation of obesity.

Although this study contributes to the measurement information on the BRFS, a formal validation study is needed. Of particular interest are the BRFS dietary and exercise behavior items, which have not been evaluated. Further research is needed to inform those using BRFS data in public health decision-making of the validity of the estimates on which such decisions are based. □

Acknowledgments

This research was supported by grant HL21906 from the National Heart, Lung and Blood Institute.

We thank Nathan Maccoby, C. Barr Taylor, and John W. Farquhar for their comments on an earlier draft of this paper, and Mary Hull for her assistance in analyzing the data.

References

1. Marks JS, Hogelin GC, Gentry EM, et al. The behavioral risk factor surveys: I. state-specific prevalence estimates of behavioral risk factors. *Am J Prev Med.* 1985;1:1–8.
2. Remington PL, Smith MY, Williamson DF, Anda RF, Gentry EM, Hogelin GC. Design, characteristics, and usefulness of state-based behavioral risk factor surveillance: 1981–87. *Public Health Rep.* 1988;103:366–375.
3. Anda RF, Dodson DL, Williamson DF, Remington PL. Health promotion data for state health departments: telephone versus in-person survey estimates of smoking and alcohol use. *Am J Health Promotion.* 1989;4:32–36.
4. Centers for Disease Control. Behavioral risk factor surveillance: selected states, 1986. *MMWR.* 1987;36:252–254.
5. Centers for Disease Control. Behavioral risk factor surveillance: selected states, 1987. *MMWR.* 1989;38:469–473.
6. Remington PL, Forman MR, Gentry EM, Marks JS, Hogelin GC, Trowbridge FL. Current smoking trends in the United States. *JAMA.* 1985;253:2975–2978.
7. Forman MR, Trowbridge FL, Gentry EM, Marks JS, Hogelin GC. Overweight adults in the United States: the behavioral risk factor surveys. *Am J Clin Nutr.* 1986; 44:410–416.
8. Bradstock MK. The behavioral risk factor surveys: III. chronic heavy alcohol use in the United States. *Am J Prev Med.* 1985; 1:15–20.
9. Shea S, Aryeh DS, Lantigua R, Basch CE. Reliability of the Behavioral Risk Factor Survey in a triethnic population. *Am J Epidemiol.* 1991;133:489–500.
10. Gentry EM, Kalsbeek WD, Hogelin GC, et al. The behavioral risk factor surveys: II. design, methods, and estimates from combined state data. *Am J Prev Med.* 1985;1:9–14.
11. Anda RF, Dodson DL, Williamson DF, Remington PL. Health promotion data for state health departments: telephone versus in-person survey estimates of smoking and alcohol use. *Am J Health Promotion.* 1989;4:32–36.
12. Farquhar JW, Fortmann SP, Maccoby N, et al. The Stanford Five-City Project: design and methods. *Am J Epidemiol.* 1985;122:323–334.
13. Waksberg J. Sampling methods for random digit dialing. *J Am Stat Assoc.* 1978;73:40–46.
14. Jackson C, Winkleby MA, Flora JA, Fortmann, SP. Utilization of educational resources for cardiovascular risk reduction in the Stanford Five-City Project. *Am J Prev Med.* 1991;7(2):(in press).
16. Fortmann SP, Marcuson RM, Bitter PH, Haskell WL. A comparison of the sphygmetrics SR-2 automatic blood pressure recorder to the mercury sphygmomanometer in population studies. *Am J Epidemiol.* 1981;114:836–844.
17. Fortmann SP, Winkleby MA, Flora JA, Haskell WL, Taylor CB. Effect of long-term community health education on blood pressure and hypertension control: the Stanford Five-City Project. *Am J Epidemiol.* 1990;132:629–645.
18. Andres R, Elahi D, Tobin JD, Muller DC, Brant L. Impact of age on weight goals. *Ann Intern Med.* 1985;103:1030–1033.
19. US Department of Health, Education and Welfare. *Lipid Research Clinics Manual of Laboratory Operations.* Vol 1. *Lipid and Lipoprotein analysis.* Washington DC: US Govt Printing Office; 1974. DHEW publication NIH 75628.
20. Cleary PD, Angel R. The analysis of relationships involving dichotomous dependent variables. *J Health Soc Behav.* 1984;25:334–348.
21. Hagen DE, Collier CM. Must respondent selection procedures for telephone surveys be invasive? *Public Opinion Q.* 1983; 47:547–556.
22. Kaplan RM, Simon HJ. Compliance in medical care: reconsideration of self-predictions. *Annals Behav Med.* 1990;12:66–71.
23. Palta M, Prineas RJ, Berman R, Hannan P. Comparison of self-reported and measured height and weight. *Am J Epidemiol.* 1982;115:223–230.
24. Stewart AW, Jackson RT, Ford MA, Beaglehole R. Underestimation of relative weight by use of self-reported height and weight. *Am J Epidemiol.* 1987;125:122–126.

Smokeless Tobacco, Cardiovascular Risk Factors, and Nicotine and Cotinine Levels in Professional Baseball Players

David Siegel, MD, MPH, Neal Benowitz, MD, Virginia L. Ernster, PhD, Deborah G. Grady, MD, MPH, and Walter W. Hauck, PhD

ABSTRACT

Background. The use of smokeless tobacco (ST), which has increased in popularity over the past 2 decades, results in considerable systemic exposure to nicotine. Nicotine might contribute to atherosclerosis by an effect on cardiovascular risk factors.

Methods. The effects of ST use on cardiovascular risk factors and cotinine and nicotine levels were studied in 1061 professional baseball players during spring training in 1988 and 1989.

Results. Of the study participants 477 (45%) were users. ST use was more common among Whites (55%) than among Blacks (29%) or Hispanics (21%), and users reported heavier consumption of alcohol ($p < .001$) and had higher mean serum caffeine levels ($p < .001$) than nonusers. ST users did not differ from nonusers in adjusted levels of systolic and diastolic blood pressure, pulse, and total or HDL-cholesterol. Among ST users, participants using snuff had higher mean serum cotinine levels than those who used chewing tobacco ($p < .001$). There was no association between serum cotinine levels and adjusted levels of any cardiovascular risk factor studied. However, higher diastolic blood pressures were associated with higher mean serum nicotine levels ($p = .02$).

Conclusions. Smokeless tobacco use has at most a modest effect on cardiovascular risk factors in young physically fit men. (*Am J Public Health.* 1992;82:417-421)

Introduction

The use of smokeless tobacco (ST)—oral snuff or chewing tobacco—increased in popularity in the 1970s and 1980s. Approximately 10 million Americans used ST in the past year; 3 million were less than 21 years of age.[1] Among young men, the use of moist snuff has become particularly widespread.[2] Smokeless tobacco use results in considerable systemic exposure to nicotine.[3,4] The results of studies of cigarette smokers have led to concern that nicotine may increase the risk of atherosclerotic vascular disease.[5] Nicotine may contribute to atherosclerosis by affecting lipid metabolism, coagulation, hemodynamic status, or all three. Smokers have higher levels of low-density lipoprotein (LDL) and very-low-density lipoprotein (VLDL) cholesterol and lower levels of high-density lipoprotein (HDL) cholesterol than do nonsmokers.[6-8] Cigarette smoking or the administration of nicotine acutely increases heart rate and blood pressure, although in habitual cigarette smokers blood pressure tends to be lower than in nonsmokers.[9] Circulating leukocytes, primarily neutrophils, are higher in smokers than in nonsmokers, and leukocyte counts return toward nonsmoking levels after smoking cessation.[10,11] Neutrophils may generate oxygen-free radicals and other substances that promote thrombosis, and they may accelerate endothelial cell growth, both of which could contribute to the acceleration of atherosclerotic vascular disease.[5,12] Whether nicotine exposure from ST use has similar effects on lipid metabolism, neutrophils, blood pressure, and heart rate, and thus predisposes to atherosclerotic vascular disease, has not been established.

In studies performed on a research ward, ST use resulted in systemic absorption of nicotine and cardiovascular effects similar to those observed with cigarette smoking.[13] In reports of small numbers of individuals, ST use has been associated with hypertension, perhaps related to ST's salt content or to nicotine-related catecholamine release.[13-15] One recent study reported an association between ST use and elevated levels of total serum cholesterol.[8]

We recently completed a comprehensive study of the health effects of ST use in professional baseball players.[16] A part of the study's purpose was to determine whether ST use was associated with changes in cardiovascular risk profiles. Specifically, we measured blood pressure, heart rate, total serum cholesterol, HDL cholesterol, and white blood cell counts. We also quantitated nicotine intake from ST by measuring plasma nicotine and its metabolite cotinine in plasma.

Methods

Study Population

A detailed discussion of the study methods and highlights of findings from

David Siegel and Deborah G. Grady are with the Department of Epidemiology and Biostatistics and the Department of Medicine; Neal Benowitz is with the Department of Medicine; and Virginia L. Ernster and Walter W. Hauck are with the Department of Epidemiology and Biostatistics—all at the University of California, San Francisco.

Requests for reprints should be sent to David Siegel, MD, Prevention Sciences Group, 74 New Montgomery, San Francisco, CA 94105.

This paper was submitted to the journal April 10, 1991, and accepted with revisions September 20, 1991.

Editor's Note: See related Commentary by Connolly on page 351.

the first year of the project have been presented elsewhere.[16] We report here detailed findings from the first 2 years of the study on cardiovascular risk factors and nicotine and cotinine levels in the participants.

During February and March of 1988 and 1989, we studied 1061 members of the seven major league professional baseball teams and their associated minor league teams that conduct spring training in the greater Phoenix and Tucson, Ariz, area. All players and coaches on the teams were invited to participate in the study.

After giving informed consent, all participants completed a questionnaire that provided demographic data and information about cigarette smoking and coffee and alcohol consumption. There were only 41 current cigarette smokers and they were excluded from analyses. Detailed information concerning patterns of ST use, including type and brand used most often, was collected. Analyses of type (snuff or chewing tobacco) and brand were based on the type and brand usually used. The amount of ST used was based on the number of cans of snuff or pouches of chewing tobacco reportedly used per week. Because many participants who usually used one type of ST occasionally used the other type, the combined number of hours that tobacco of either type was held in the mouth per day was also calculated. Recency of use was determined as hours since ST was last used.

The participants were classified on the basis of their self-reported ST use as nonusers (those who had never used ST or who had used ST in the past but never more frequently than once a month) and users (those who had used ST within the past week). Former users (those who had used ST more than once a month in the past but had not used ST within the previous month) and infrequent current users (those who had used ST within the past month but not within the past week) were excluded from analyses. For any player for whom we had data from both 1988 and 1989, we used only the data from the year for which the data were most complete. If data from both years were equally complete, 1989 data were used for that player. Data from a total of 477 ST users and 584 nonusers are included in this report.

Blood Pressure and Pulse

After the participants completed the questionnaire, their heart rate and blood pressure were measured twice, several minutes apart, in a sitting position. Systolic blood pressure was recorded at the level when phase 1 Korotkoff sounds were first heard and diastolic blood pressure was recorded at the beginning of phase 5 with the disappearance of sounds. To avoid the effects of physical activity, we included in our analysis of blood pressure and pulse rates only measurements from individuals who had not come to the examination directly from the playing field.

Biochemical Measurements

Blood was drawn to measure total serum cholesterol; HDL cholesterol; nicotine, cotinine, and caffeine levels; and white blood cell count. Biochemical validation of self-reported ST use included measurement of cotinine (the primary metabolite of nicotine) and thiocyanate (a marker of tobacco smoke exposure) by gas chromatography.[17] Low serum cotinine level (<12 ng/mL) together with normal serum thiocyanate levels (<85 mmol/L) was considered biochemical evidence of nonuse of tobacco.[18] This standard, applied in an initial analysis of participants from whom blood was obtained in 1988 and who reported that they neither used ST nor smoked cigarettes, indicated that 95% (357 of 376 for whom blood samples were obtained in 1988) were correctly classified biochemically as nonusers. Given the high accuracy of use status based on self-report and the fact that blood samples were not obtained from all players presented here, the results shown in Tables 1, 2, and 4 are based on self-report. Serum caffeine was measured simultaneously with nicotine and cotinine; 7-ethyl theophylline, 5-methylnicotine, and 1-methyl-5-(2-pyridyl)-pyrrolidine, respectively, were used as internal standards.

Statistical Analysis

Estimated odds ratios were calculated and adjusted for covariates by means of multiple logistic analyses.[19] All logistic analyses were done with the CATMOD procedure on PC-SAS (SAS Institute, Cary, NC) with confidence intervals based on standard errors of the log odds ratios. Significance levels for analyses of covariance were determined with PC-SAS's GLM procedure. The means reported are least squares means, that is, estimates of the means for each use group adjusting for differences between groups in the covariates. For categorical covariates we created the appropriate indicator variables so that the least squares means are distributed around the overall mean, corresponding to the usual method of adjusting for confounding.

Results

Demographic Characteristics

The distribution of current ST users and nonusers by self-reported age, race, educational level, and alcohol consumption is presented in Table 1. Most participants (77%) in both groups were between 20 and 29 years of age and were white (67%). Participants were relatively well educated—94% had graduated from high school and 74% had completed at least some college.

Twenty-four percent reported no alcohol consumption and 7% reported an average of more than two drinks per day. Smokeless tobacco use varied by age; current-week use was reported by 26% of participants younger than 20 years, 49% of those aged 20 to 29 years, and 38% of those aged 30 years and older. Whites were more likely to report ST use during the past week (55%) than Blacks (29%) or Hispanics (21%). Current-week ST users were heavier consumers of alcohol than were nonusers ($P < .001$), and had higher mean serum caffeine levels (512 ng/ml versus 331 ng/ml, data not shown; $P < .001$) when adjusted for age and race.

Of the ST users, 75% used primarily oral snuff, the others chewing tobacco. Among current users of snuff, 70% used Copenhagen, 21% Skoal, and 6% Hawken. Of the chewers, 48% used Levi-Garett and 42% used Redman (data not shown).

Cardiovascular Risk Factors

There were no significant differences between ST users and nonusers in systolic and diastolic blood pressure, pulse, and total or HDL cholesterol levels when adjusted for age, race, alcohol use, and serum caffeine (Table 2). ST users had lower mean white blood cell counts than nonusers ($P < .01$). There were also no significant differences among ST users in systolic and diastolic blood pressure, pulse rate, white blood cell count, and total or HDL cholesterol levels when evaluated by different measures of use—years of use, hours of use per day, and hours since last use (data not shown).

Cotinine, Nicotine, and Caffeine Levels

Serum cotinine measurements were obtained in 230 of the current users. Table 3 presents cotinine levels in relation to type and patterns of ST use. Participants who used snuff had higher adjusted mean serum cotinine than participants who used

chewing tobacco ($P < .001$). Only 1% of the participants who used chewing tobacco, compared with 24% of those who used snuff, had serum cotinine levels of ≥250 ng/mL (data not shown). Participants who reported using ST more hours per day, as well as those who reported more recent use, also had higher adjusted mean serum cotinine levels ($P < .001$). Similar statistically significant results were found for adjusted mean serum nicotine levels (data not shown). Participants who had used ST for more years had higher adjusted mean serum caffeine levels than those who had used ST for a shorter period ($P = .03$, data not shown).

Cotinine and Nicotine Levels and Cardiovascular Risk Factors

There was no difference in adjusted systolic and diastolic blood pressure, pulse rate, white blood cell count, and total or HDL cholesterol levels between participants who used snuff exclusively and those who used chewing tobacco exclusively (Table 4). We examined the association between the above measures and serum cotinine and serum nicotine levels, adjusting for age, race, alcohol use, and serum caffeine (data not shown). There was no association between serum cotinine levels and any of these measures. Interestingly, we failed to confirm our previous finding of a weak inverse association between serum cotinine levels and serum HDL cholesterol levels.[16] However, higher mean serum nicotine levels were associated with higher diastolic blood pressure levels ($P = .02$), and there was a trend toward higher pulse rates ($P = .09$), total cholesterol levels ($P = .10$), and white blood cell counts ($P = .15$) with higher mean serum nicotine levels.

TABLE 1—Characteristics of Study Sample According to Smokeless Tobacco Use

	Nonuser, % (n)	User, % (n)	Total, % (n)
Age, y			
<20	13.0 (78)	5.7 (27)	10.0 (105)
20–24	49.7 (288)	56.9 (269)	52.9 (557)
25–29	21.9 (127)	26.0 (123)	23.7 (250)
30–34	7.8 (45)	7.0 (33)	7.4 (78)
35–39	2.1 (12)	1.9 (9)	2.0 (21)
>40	5.2 (30)	2.5 (12)	4.0 (42)
Race			
White	54.6 (319)	81.6 (389)	66.7 (708)
Hispanic	25.2 (147)	8.2 (39)	18.0 (186)
Black	18.0 (108)	9.4 (45)	14.4 (153)
Asian/Other	1.7 (10)	0.8 (4)	1.1 (14)
Educational level			
Elementary school	2.9 (16)	0.6 (3)	1.8 (19)
Some high school	5.2 (29)	3.4 (16)	4.4 (45)
High school graduate	22.6 (127)	16.3 (76)	19.7 (203)
Some college	46.4 (261)	58.0 (268)	51.0 (529)
College graduate	23.0 (129)	22.1 (103)	22.6 (232)
Alcohol consumption*			
Nondrinker	34.7 (200)	9.9 (47)	24.0 (247)
1–13 drinks per week	61.8 (356)	78.0 (372)	69.3 (728)
14–20 drinks per week	1.2 (7)	5.3 (25)	3.0 (32)
>20 drinks per week	2.3 (13)	6.3 (30)	4.1 (43)

Note. Total n was 1061; because not every participant answered every question on the questionnaire, the numbers in each category will not add to the total.
*$P < .001$ for trend users versus nonusers.

TABLE 2—Adjusted[a] Mean Levels of Systolic and Diastolic Blood Pressure, Pulse Rate, Total Cholesterol, HDL Cholesterol, and White Blood Cell Count by Smokeless Tobacco Use

	Nonuser		User		95% CI for Difference
	Mean	n	Mean	n	
Blood pressure (mm Hg)					
Systolic	117.1	176	117.1	127	−2.48, 2.53
Diastolic	72.1	176	71.5	127	−1.62, 2.79
Pulse (beats/min)	65.6	175	65.4	126	−2.67, 3.13
Total cholesterol (mmol/L)	4.42	489	4.39	396	−0.09, 0.15
HDL cholesterol (mmol/L)	1.30	485	1.31	395	−0.05, 0.04
WBCs × 10^9/L	6.6	419	6.2*	332	0.12, 0.64

Note. HDL indicates high density lipoprotein; WBC, white blood cells; CI, confidence interval. CIs for users minus nonusers.
[a]Adjusted for age, race, alcohol use, and serum caffeine level.
*$P < .01$ compared with nonusers.

Discussion

This group of professional baseball players is the largest population of ST users in which nicotine and associated cardiovascular risk factors have been examined. The large sample size and the high prevalence of ST use provided ample power to detect relatively minor effects of ST use on the cardiovascular risk factors studied. Despite this, we did not find differences in systolic and diastolic blood pressure, total and HDL cholesterol, and white blood cell count between ST users and nonusers. These findings are similar to those that we reported after the first year of the study[16] and contrast with the results of experimental studies of ST users in which ST use throughout the day produced the same cardiovascular effects as cigarette smoking.[13] Additionally, we analyzed the effect of ST use on serum nicotine and cotinine levels and their association with cardiovascular risk factors.

Cotinine is the major metabolite of nicotine and has a much longer half-life than nicotine.[12] For these reasons cotinine is widely used as a biochemical marker of average daily intake of nicotine. Determination of serum cotinine, however, is a relatively insensitive way to detect exactly when tobacco was last used. The mean serum cotinine level was 144 ng/mL for snuff users and 82 ng/mL for chewers. The average cotinine levels for snuff users were similar to those previously reported in a group of college athletes, but were considerably lower than those measured in a research ward study of ST use and cigarette smoking.[3,13] These cotinine levels are also lower than the average level of 300 ng/mL found in cigarette smokers.[12] Since a cotinine level of 100 ng/mL corresponds to a nicotine intake of about 12 mg,[12] we estimated that daily intake of nicotine was 17.0 mg for snuff users and 9.9 mg for chewing tobacco users. The low

TABLE 3—Adjusted[a] Mean Levels of Serum Cotinine in Smokeless Tobacco Users by Type of Smokeless Tobacco Use, Years of Use, Hours of Use Per Day, and Hours Since Last Use

	n	Mean Serum Cotinine (ng/ml) (SE)
Type of smokeless tobacco*		
Snuff	182	143.9 (7.4)
Chewing Tobacco	48	82.1 (15.1)
Years of use		
3 or less	31	91.1 (18.1)
4–6	97	145.6 (10.2)
7–9	54	123.9 (13.9)
10 or more	48	135.2 (14.6)
Hours of use per day*		
0.0–0.5	73	104.3 (12.2)
0.5–1.0	49	131.7 (14.2)
1.0–1.5	33	134.8 (17.6)
1.5–2.0	22	120.8 (21.4)
2.0–4.0	38	165.3 (16.3)
>4.0	15	178.1 (26.5)
Hours since last use*		
0–1	61	174.0 (13.1)
1–12	60	185.6 (13.2)
12–24	78	87.9 (11.4)
>24	31	49.1 (19.2)

[a]Each analysis is adjusted for other variables in the table.
*$P < .001$; for the two "hours" variables, this P is from trend tests using the continuous form of the variables.

TABLE 4—Adjusted[a] Mean Levels of Systolic and Diastolic Blood Pressure, Pulse Rate, Total and HDL Cholesterol, and White Blood Cell Count by Type of Smokeless Tobacco Use[b]

	Snuff		Chewing Tobacco		95% CI for difference
	Mean	n	Mean	n	
Blood Pressure					
Systolic	115.3	69	119.3	26	−9.10, 1.07
Diastolic	71.9	69	70.9	26	−2.19, 5.14
Pulse	64.5	69	65.4	25	−4.22, 6.05
Total Cholesterol	4.34	180	4.39	48	−0.33, 0.21
HDL-Cholesterol	1.33	179	1.33	48	−0.11, 0.09
White Blood Cell Count	6.1	154	6.2	33	−0.66, 0.49

Note. Same abbreviations and units as in Table 2.
[a]Adjusted for age, race, alcohol use, serum caffeine level, hours of smokeless tobacco use per day, time since last ST use and years of ST use.
[b]Subjects included in this table are exclusive users of either snuff or chewing tobacco.

cotinine levels in our subjects were consistent with relatively light use of ST, as reflected by an average of 1.1 hours of snuff use per day and 0.7 hours of chewing tobacco use per day. As expected, serum cotinine levels tended to be higher with greater duration of ST use and greater number of hours of ST use per day, and was lower with greater number of hours since the last dose of ST. Overall, the serum cotinine data indicate that baseball players are relatively light users of ST compared with other populations. Our data also indicate that ST use in baseball players is seasonal; about 50% of the baseball players in our study who reported current-week ST use indicated that they use ST primarily during the baseball season and rarely use it during the off season. The pattern of ST use and low levels of nicotine intake suggests that baseball players may be less dependent on nicotine than are other tobacco users.

Our finding that cotinine levels were lower with chewing tobacco use than with snuff use was surprising in light of the observation that nicotine intake for a single chew is substantially greater than that from a single use of snuff, and daily nicotine levels with ad lib use of both in an experimental study have been found to be similar.[4,13] Our results suggest that chewers use ST in a manner that does not extract as much nicotine from ST as snuff users. The relationship between hours of use per day and serum cotinine was not highly correlated, suggesting that nicotine intake may be more efficient with shorter periods of daily use than with longer periods of use. This type of dose response could be observed if ST is held in the mouth for long periods of time, since the most rapid absorption of nicotine through ST occurs in the first 15 to 20 minutes.[4]

We found an association between higher serum nicotine levels and increased diastolic blood pressure and trends between higher plasma nicotine levels and increased heart rate and white blood cell count. Serum nicotine reflects the recency of tobacco use, and changes in blood pressure, pulse rate, and white blood cell count might reflect a direct pharmacologic response.[12] Our diastolic blood pressure finding is consistent with findings from other studies showing that ST use acutely increases blood pressure.[20,21]

The smokeless tobacco users in this study drank more alcohol than did nonusers. This finding is similar to those of other studies of ST use, in both adolescents and adults.[22] We also found that plasma caffeine levels were higher in ST users than nonusers, perhaps reflecting a tendency toward more psychoactive drug use in ST users. Other studies have found that cigarette smokers drink more caffeine-containing beverages than do nonsmokers.[23] Thus, future studies of the cardiovascular effects of ST should control for the potentially confounding effects of caffeine and alcohol.

Although our findings did not show major adverse effects of ST use on cardiovascular risk factors, we cannot exclude the possibility that daily ST use for many years may have an adverse effect on the cardiovascular system. Previous reports have described long-term harmful effects of ST use, including a strong association with oral cancer.[24–26] Our participants differed from those in other studies of ST use in that their ST use was generally seasonal and of short duration (less than 10 years). Our participants were also professional athletes who exercised regularly and who were generally physically fit, factors that might minimize the harmful effects of ST on the cardiovascular system. Similar large studies should be done in older populations of more sedentary in-

dividuals to detect the impact of ST use in individuals more prone to cardiac disease. □

Acknowledgment

This research was supported by grants 5 PO1 DE08547-02 from the National Institute of Dental Research and DA 02277 from the National Institute on Drug Abuse, Bethesda, Md.

References

1. Consensus Conference. Health applications of smokeless tobacco. *JAMA.* 1986;255:1045–1048.
2. *Reducing the Health Consequences of Smoking: 25 Years of Progress: A Report of the Surgeon General.* Rockville, Md: Office on Smoking and Health; 1989.
3. Gritz ER, Baer-Weiss V, Benowitz NL, Van Vunakis H, Jarvik ME. Plasma nicotine and cotinine concentrations in habitual smokeless tobacco users. *Clin Pharmacol Ther.* 1981;30:201–209.
4. Benowitz NL, Porchet H, Sheiner L, Jacob P. Nicotine absorption and cardiovascular effects with smokeless tobacco use: comparison with cigarettes and nicotine gum. *Clin Pharmacol Ther.* 1988;44:23–28.
5. Benowitz NL. Nicotine and cardiovascular disease. In: Adlhofer F, Thurau K, eds. *Effects of Nicotine on Biological Systems.* Basel, Switzerland: Birkhauser Verlag; 1991:579–596.
6. Brischetto CS, Connor WE, Connor SL, Matarazzo JD. Plasma lipid and lipoprotein profiles of cigarette smokers from randomly selected families: enhancement of hyperlipidemia and depression of high density lipoprotein. *Am J Cardiol.* 1983;52: 675–680.
7. Dwyer JH, Rieger-Ndakorerwa GE, Semmer NU, Fuchs R, Lippert P. Low-level cigarette smoking and longitudinal change in serum cholesterol among adolescents. *JAMA.* 1988;259:2857–2862.
8. Tucker LA. Use of smokeless tobacco, cigarette smoking, and hypercholesterolemia. *Am J Public Health.* 1989;79:1048–1050.
9. Benowitz NL, Sharp DS. Inverse relation between serum cotinine concentration and blood pressure in cigarette smokers. *Circulation.* 1989;80:1309–1312.
10. Grimm RH, Neaton JD, Ludwig W. Prognostic importance of the white blood cell count for coronary, cancer, and all-cause mortality. *JAMA.* 1985;254:1932–1937.
11. Knoke JD, Hunninghoke DB, Heiss G. Physiological markers of smoking and their relation to coronary heart disease. *Arteriosclerosis.* 1987;7:477–482.
12. Benowitz NL, Porchet H, Jacob P III. Pharmacokinetics, metabolism, and pharmacodynamics of nicotine. In: Wonnacott S, Russell MAH, Stolerman, IP, eds. *Nicotine Psychopharmacology—Molecular, Cellular and Behavioral Aspects.* New York, NY: Oxford University Press; 1990: 112–157.
13. Benowitz NL, Jacob P III, Yu L. Daily use of smokeless tobacco: systemic effects. *Ann Intern Med.* 1989;111:112–116.
14. Benowitz NL. Pharmacologic aspects of cigarette smoking and nicotine addiction. *N Engl J Med.* 1988;319:1318–1330.
15. Hampson NB. Smokeless is not saltless. *N Engl J Med.* 1985;312:919–920. Letter.
16. Ernster VL, Grady DG, Greene JC, et al. Smokeless tobacco use and health effects among baseball players. *JAMA.* 1990;264: 218–224.
17. Jacob P III, Wilson M, Benowitz NL. Improved gas chromatography method for the determination of nicotine and cotinine in biologic fluids. *J Chromatogr.* 1981;222: 61–70.
18. Cummings SR, Richard RJ. Optimum cutoff points for biochemical validation of smoking status. *Am J Public Health.* 1988;78:574–575.
19. Hosmer DW, Lemeshow S. *Applied Logistic Regression.* New York, NY: John Wiley & Sons Inc; 1989.
20. Squires WG Jr, Brandon TA, Zinkgraf S, et al. Hemodynamic effects of oral smokeless tobacco in dogs and young adults. *Prev Med.* 1984;13:195–206.
21. Wells DG, Rustick JM. Hypertension from smokeless tobacco. *Anesthesiology.* 1986;65:339. Letter.
22. Gritz ER, Kassir C, McCarthy WJ. Smokeless tobacco use in the United States: present and future trends. *Ann Behavioral Med.* 1985;7:24–27.
23. Istran J, Matarazzo JD. Tobacco, alcohol and caffeine use: a review of their interrelationships. *Am Psychol Bull.* 1984;95:301–326.
24. *The Health Consequences of Using Smokeless Tobacco: A Report of the Advisory Committee to the Surgeon General.* Bethesda, Md: National Institutes of Health; April, 1986. Publication N086-2874.
25. *Evaluation of the Carcinogenic Risk of Chemicals to Humans.* Lyon, France: International Agency for Research on Cancer; 1985;37.
26. Winn DM, Blot WJ, Shy CM, Pickle LW, Toledo A, Fraumeni JF Jr. Snuff dipping and oral cancer among women in the southern United States. *N Engl J Med.* 1981; 304:745–749.

Public Health Briefs

Morbidity among Pediatric Motor Vehicle Crash Victims: The Effectiveness of Seat Belts

ABSTRACT

It is well established that seat belts reduce mortality and morbidity among children. Data are presented for 413 children injured severely enough in motor vehicle crashes to require hospitalization. Of the unrestrained children, 4.5% died, compared with 2.4% of the belted children. Unrestrained children had a higher proportion of injuries in four of five anatomical regions, were more severely injured, stayed longer in the hospital, and were 15% more likely than belted children to be discharged with impairments. (*Am J Public Health.* 1992;82:422–425)

J. Scott Osberg, MA, and Carla Di Scala, PhD

Introduction

Despite efforts to reduce motor vehicle fatalities and injuries via restraint laws, speed limits, and alcohol awareness programs, these injuries remain one of the most frequent causes of death and disability for children.[1] Although there is little doubt that use of seat belts reduces fatalities,[2,3] serious injuries, and concomitant medical costs,[3,4–11] belts can cause certain injuries, particularly in the 4-to-9 age group, in which children are too big for child safety seats and too small for adult belts.[12] Injuries associated with belt use include chest and abdominal injuries,[13–16] cervical spine fractures,[17] cardiac contusion,[18] and rupture of the stomach.[19]

This paper focuses on the effectiveness of seat belts in reducing injuries among survivors aged 4 to 14 who are admitted to hospitals. The study contrasts outcomes for belted versus unrestrained children among the middle range of children involved in motor vehicle crashes—those requiring hospitalization.

Methods

Data are from the National Pediatric Trauma Registry, Phase II.[20] The Phase II registry population contained 7156 children injured from all causes from October 1988 to March 1990, 1150 of whom were injured in motor vehicle crashes. Children in crashes not occurring on the road, those riding in pickup trucks or buses, children younger than 4 or older than 14 years, and 43 cases in which restraint use was unknown were excluded. Uninjured children, those receiving only emergency room services, and children dead at the scene were excluded. Of the remaining 413 children, 290 were unrestrained and 123 were belted. This hospital-based sample was expected to reduce differences between the two groups, as children on either end of the motor vehicle crash spectrum (those unscathed and those who died at the scene) were excluded. No information was available on vehicular speed(s) or side of impact, which may have influenced outcome.

Outcome variables include mortality, severity of injury, body regions affected, impairments, and utilization of health services. Three injury severity measures are considered: the Abbreviated Injury Scale (AIS),[21] the Injury Severity Score (ISS),[22] and the Glasgow Coma Scale.[23] Injured children were classified according to the most severely involved ISS body region, identified by the maximum AIS score.[24] The ranking system of MacKenzie et al.[24] is used to assign the principal injury when there are ties for the most severely injured AIS body region. Impairments at discharge were assessed in nine areas of daily functioning: vision, hearing, speech, self-feeding, bathing, dressing, walking, cognition, and behavior.

Results

Of the 123 belted children, 40.5% were male; 50.2% of the 290 unrestrained

The authors are with the Department of Rehabilitation Medicine, Rehabilitation Research and Training Center in Rehabilitation and Childhood Trauma, Tufts-New England Medical Center Hospitals, Boston, Mass. J. Scott Osberg is also with the Tufts University School of Medicine.

Requests for reprints should be sent to J. Scott Osberg, MA, Department of Rehabilitation Medicine, Rehabilitation Research and Training Center in Rehabilitation and Childhood Trauma, Tufts-New England Medical Center Hospitals, 750 Washington Street, NEMC 75 K/R, Boston, MA 02111.

This paper was submitted to the Journal January 14, 1991, and accepted with revisions June 10, 1991.

children were male. The average age of the children in both the belted and unrestrained groups was 8.5 years. The two groups were similar in proportions of front- and back-seat passengers and crash types (Table 1). However, unrestrained children tended to be more severely injured, to have higher mortality rates (Table 2), and to use more services (Table 3). As expected, unrestrained children were more likely to have head/neck injuries, as well as injuries to the face, thorax, and extremities (Figure 1). As found in earlier studies,[14–16] belted children were somewhat more likely to have injuries to the abdomen. Head/neck and thoracic injuries posed the greatest survival threat to both groups (Figure 2).

The abdominal region deserves special attention, as it is the one area where belted children were more severely injured. Whereas 22.3% of the unrestrained children had moderate to severe abdominal AIS scores, 31.7% of belted children had AIS scores in this range (Figure 2). Table 4 displays the constellation of injured body regions for the subset of children with primary abdominal injuries. Belted children had either abdominal injuries alone (57.1%) or abdominal and external injuries (42.9%), whereas 33.3% of the unrestrained children had injuries to a variety of body regions.

Unrestrained children were more likely to have head/neck, extremity, or thorax injury than were belted children, who were more likely to have abdominal or "other" as their primary injury (Figure 3). As determined by the ISS of the primary injury, unrestrained children, particularly those classified with primary injuries to the head/neck or abdomen, were more seriously injured. Finally, 70.6% of the belted children, compared with 54.8% of the unrestrained children, were discharged with no impairments (Figure 4).

TABLE 1—Crash-Related Factors

	No Restraint		Seat Belt	
	No.	%	No.	%
Occupant ($P > .05$)				
Front	110	37.9	50	40.7
Back	138	47.6	54	43.9
Unspecified	42	14.4	19	15.4
Total	290	99.9	123	100.0
Type of crash ($P > .05$)				
Collision with another motor vehicle	226	77.9	104	84.6
Collision with object (tree, telephone pole, etc.)	43	14.8	11	8.9
Loss of control (no collision)	20	6.9	8	6.5
Other	1	0.3	0	0.0
Total	290	99.9	123	100.0

Note. Because of rounding error, some percentages do not add to 100.

TABLE 2—Morbidity and Mortality

	No Restraint		Seat Belt	
	No.	%	No.	%
Injury Severity Score ($P > .05$)				
1–9	170	58.6	87	70.7
10–19	70	24.1	20	16.3
20+ (severe)	50	17.2	16	13.0
Total	290	99.9	123	100.0
Glascow Coma Scale - Among Head Injuries only ($P > .05$)				
13–15	110	67.5	39	72.2
9–12	16	9.8	7	13.0
3–8 (severe)	37	22.7	8	14.8
Total	163	100.0	54	100.0
Mortality ($P > .05$)				
Alive	277	95.5	120	97.6
Dead	13	4.5	3	2.4
Total	290	100.0	123	100.0

Note. Because of rounding error, some percentages do not add to 100.

Discussion

Analysis of this multi-institutional sample demonstrates that seat belts do more than "merely" reduce the incidence of injury and concomitant mortality. Despite the homogenizing effect of this hospital-based sample, the results suggest that belts reduce morbidity, with everything this reduction implies in terms of quality of life and reduced health care costs. Belted survivors were less severely injured, were hospitalized for shorter periods, were less likely to be discharged with residual impairments, and consequently were less likely to need health services following discharge. Thus, not only do seat belts reduce health costs by preventing injuries,[4–6,10] but even in the subset of children who require hospitalization, the use of seat belts results in the consumption of fewer resources.

Belted children were found to have more abdominal injuries of greater severity than did unrestrained children. However, unrestrained children with primary abdominal injuries also had injuries to other body regions, which left them more severely involved overall. Although belted children had more severe abdominal injuries, the low risk of sustaining serious belt-related abdominal injuries is overshadowed by the greater risk of sustaining other serious injuries if seat belts are not used.[14] Although clinicians should be alert to signs of "seat belt syndrome" among belted children,[14–16,25] we agree with Cameron[26] that seat belt syndrome should be de-emphasized so parents don't misinterpret the findings and allow children to go unbelted.

The optimal protection system has not yet been designed; continued improvements are needed.[14,26–28] Reid et al.[29] support Agran et al.[30] in recommending the "use of specially designed passenger restraints in children up to the age of 10 years."[29] However, Reid et al. also report

TABLE 3—Utilization of Health Services

	No Restraint		Seat Belt	
Trauma center days ($P > .05$)				
Mean	7.37		6.34	
Median	3.00		2.00	
Standard error	0.79		1.33	
Minimum	0.00		0.00	
Maximum	133.00		140.00	
ICU days ($P < .05$)				
Mean	1.39		0.83	
Median	0.00		0.00	
Standard error	0.18		0.16	
Minimum	0.00		0.00	
Maximum	28.00		12.00	
Treated in Operating Room ($P > .05$)				
Yes, no. (%)	81	(27.9)	33	(26.8)
No, no. (%)	209	(72.1)	90	(73.2)
Total, no. (%)	290	(100.0)	123	(100.0)
Discharge location ($P > .05$)				
Home, no. (%)	246	(84.8)	114	(92.7)
Hospital, no. (%)	8	(2.8)	2	(1.6)
Rehabilitation, no. (%)	12	(4.1)	2	(1.6)
Other, no. (%)	24	(8.3)	5	(4.1)
Total, no. (%)	290	(100.0)	123	(100.0)

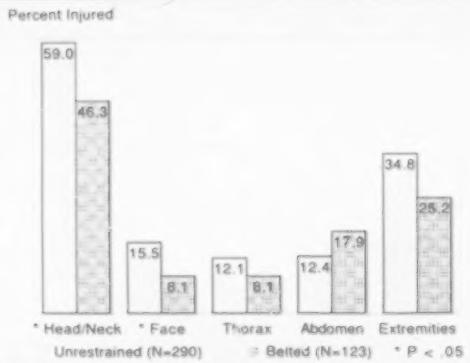


FIGURE 1—Percentages of unrestrained and belted children with injured body areas. (Because of multiple injuries suffered by some children, percentages for unrestrained and belted children each add to more than 100.)

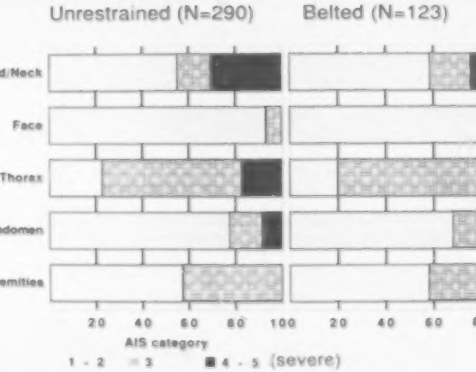


FIGURE 2—Abbreviated Injury Scale (AIS) categories among children injured in five body areas (unrestrained: n = 290; belted: n = 123).

that seat belts are associated with pediatric chance fractures: "While seatbelts may in fact save lives and reduce mortality, their use may inadvertently lead to an increased morbidity of the survivors of high-impact motor vehicle accidents in the pediatric population."[29] This speculation is not supported by the data in our sample, where there appears to be no increased morbidity among belted survivors. In fact, the reductions in the incidence of injury and in mortality rates for belted children noted elsewhere[3,4–11] were matched in our sample by reductions in short- and long-term morbidity. Seat belts are an effective means of prevention, and even current adult designs are capable of producing important reductions in morbidity in pediatric populations. In short, any restraint is better than no restraint.[27] □

Acknowledgments

This work of the Medical Rehabilitation Research and Training Center in Rehabilitation and Childhood Trauma was supported in part by research grant H133B80009 from the National Institute of Disability and Rehabilitation Research, US Department of Education, Washington, DC, and by a Centers for Disease Control grant to the Harvard Injury Control Center, Harvard School of Public Health, Boston, Mass.

The authors wish to thank the physicians and the individual data collectors at the 50 participating National Pediatric Trauma Centers. In addition, we thank Stephen M. Haley for helpful comments on earlier drafts of the manuscript. We also wish to acknowledge Laura J. Chin and Cheryl Grant for their contributions to the National Pediatric Trauma Registry.

References

1. Tepas JJ, DiScala C, Ramenofsky ML, Barlow B. Mortality and head injury: the pediatric perspective. *J Pediatr Surg.* 1990;25:92–96.
2. Evans L. Rear seat restraint system effectiveness in preventing fatalities. *Accid Anal Prev.* 1986;18:229–241.
3. Sewell CM, Hull HF, Fenner JF, Graff H, Pine J. Child restraint law effects on motor vehicle accident fatalities and injuries: the New Mexico experience. *Pediatrics.* 1986;78:1079–1084.
4. National Committee for Injury Prevention and Control. Injury prevention: meeting the challenge. *Am J Prev Med.* 1989; 5(suppl):279–280.
5. Barancik JI, Kramer CF, Thode HC, Harris D. Efficacy of the New York State Seat Belt Law. Preliminary Assessment of Occurrence and Severity. *Bull N Y Acad Med.* 1988;64:742–749.
6. Rice DP, Kaufman SR, McLaughlin E. Cost of injury—United States: a report to Congress. *JAMA.* 1989;262:2803–2804.
7. Christian MS, Bullimore DW. Reduction in accident injury severity in rear seat passengers using restraints. *Injury.* 1989;20:262–264.
8. Chorba TL, Reinfurt D, Hulka BS. Effi-

TABLE 4—Constellation of Injuries among Children with Primary Injuries of the Abdomen

Head	Face	Thorax	Abdomen	Extremity	External	No Belt (n = 12) %	Belt (n = 14) %
		X	X	X	X	8.3	
	X		X	X	X	8.3	
X			X	X		8.3	
	X		X		X	8.3	
			X		X	33.3	42.9
			X			33.3	57.1
						99.8	100.0

Note. Because of rounding error, some percentages do not add to 100.

cacy of mandatory seat-belt use legislation: The North Carolina Experience from 1983 through 1987. *JAMA.* 1988;260:3593–3597.

9. Decker MD, Dewey MJ, Hutcheson RH, Shaffner W. The use and efficacy of child restraint devices. The Tennessee Experience. *JAMA.* 1984;252:2571–2575.
10. Orsay EM, Turnbull TL, Dunne M, Barrett JA, Langenberg P, Orsay CP. Prospective study of the effect of safety belts on morbidity and health care costs in motor vehicle accidents. *JAMA.* 1988;260:3598–3603.
11. Graham JD. Preventing Automobile Injury: New Findings from Evaluation Research. Dover, Mass: Auburn House Publishing Co; 1988.
12. Agran PF, Winn D, Dunkle D. Injuries among 4 to 9 year old restrained motor vehicle occupants by seat location and crash impact site. *Am J Diseases Child.* 1989;143:1317–1321.
13. Arajarvi E, Santavirta S. Chest injuries sustained in severe traffic accidents by seat-belt wearers. *J Trauma.* 1989;29:37–41.
14. Asbun HJ, Irani H, Roe EJ, Bloch JH. Intra-abdominal seatbelt injury. *J Trauma.* 1990;30:189–193.
15. Appleby JP, Nagy AG. Abdominal injuries associated with the use of seatbelts. *Am J Surg.* 1989;157:457–458.
16. Denis R, Allard M, Atlas H, Farkouh E. Changing trends with abdominal injury in seatbelt wearers. *J Trauma.* 1983;23:1007–1008.
17. Fuchs S, Barthel MJ, Flannery AM, Christoffel KK. Cervical spine fractures sustained by young children in forward-facing car seats. *Pediatrics.* 1989;84:348–354.
18. Muwanga CL, Cole RP, Sloan JP, Bruce E, Dove AF, Dave SH. Cardiac contusion in patients wearing seat belts. *Injury.* 1986;17:37–39.
19. Baker R, Perry EP, Fossard DP. Traumatic rupture of the stomach due to a seat belt. *Injury.* 1986;17:47–48.
20. Tepas JJ, Ramenofsky ML, Barlow B, et al. National pediatric trauma registry. *J Pediatr Surg.* 1989;24:156–158.
21. Gennarelli TA, Baker SP, Bryant RW, et al. *The Abbreviated Injury Scale 1985 Revision.* Arlington Heights, Ill: American Association for Automotive Medicine; 1985.
22. Baker SP, O'Neill B, Haddon W, Long W. The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma.* 1974;14:187–196.
23. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: a practical scale. *Lancet.* 1974;2:81.
24. MacKenzie EJ, Siegel JH, Shapiro S, Moody M, Smith RT. Functional recovery and medical costs of trauma: an analysis by type and severity of injury. *J Trauma.* 1988;28:281–297.
25. Garrett JW, Brunstein PW. The seat belt syndrome. *J Trauma.* 1962;2:220.
26. Cameron BH. Injuries in children wearing seat belts. *Can Med Assoc J.* 1986;134:308–309.
27. Stylianos S, Harris BH. Seatbelt use and patterns of central nervous system injury in children. *Pediatr Emerg Care.* 1990;6:4–5.
28. Wagenaar AC, Margolis LH. Effects of a mandatory safety belt law on hospital admissions. *Accid Anal Prev.* 1990;22:253–261.
29. Reid AB, Letts RM, Black GB. Pediatric chance fractures: association with intra-abdominal injuries and seatbelt use. *J Trauma.* 1990;30:384–391.
30. Agran PF, Dunkle DE, Winn DG. Injuries to a sample of seatbelted children evaluated and treated in a hospital emergency room. *J Trauma.* 1987;27:58–64.

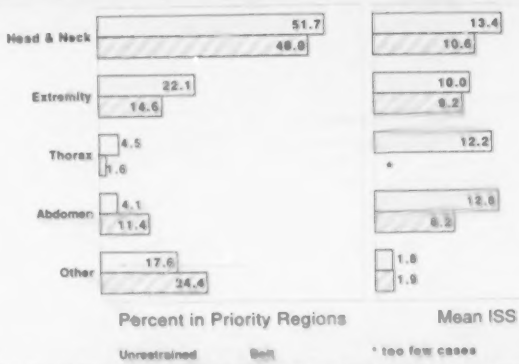


FIGURE 3—Percentages in injury priority regions and mean Injury Severity Score (ISS). Each child was assigned only one injury priority region. The ISS is a measure of the overall severity in all regions; high scores are associated with more severe injuries.

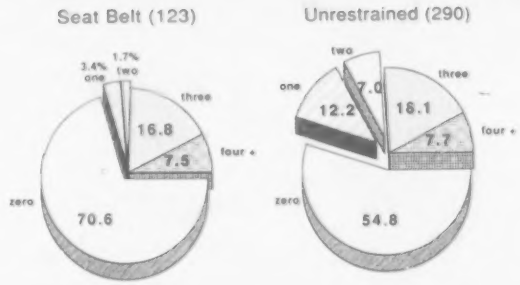


FIGURE 4—Number of impairments at discharge for belted and unrestrained children.

Pediatric Gonococcal Infection, Florida, 1984 to 1988

Jean-Claude A. Desenclos, MD, David Garrity, and Jack Wroten

ABSTRACT

We reviewed Florida pediatric gonococcal (GC) infection surveillance data collected between 1984 and 1988. The incidence rate was 11 per 100 000 per year for the age group 1 month through 9 years. Incidence rates were higher for females, other-than-Whites, and residents of rural counties than for males, Whites, and residents of urban counties. The ratio of pediatric GC cases in children younger than 10 years per 1000 adult male GC cases, a substitute measure for the proportion of males that may have perpetrated child sexual abuse, was 3.3 for Whites and 4.9 for other-than-Whites, and was higher for residents of nonmetropolitan counties (9.7) than for residents of metropolitan counties (4.2). These data highlight the importance of GC infection in children and suggest that routine surveillance of pediatric GC infection may be a useful tool for monitoring the occurrence of child sexual abuse. (*Am J Public Health*. 1992;82:426–428)

Introduction

Sexual abuse of children is receiving increasing attention; there are an estimated 100 000 to 500 000 victims per year in the United States.[1] Although misidentification of organisms such as *Neisseria gonorrhoeae* has been reported in children,[2] the isolation of a sexually transmissible pathogen from a child, when infection in utero or during delivery can be excluded, is presumed to be evidence of sexual abuse.[3] Gonorrhea is the most frequently reported sexually transmitted disease (STD) in children who have been sexually abused.[4] To estimate the public health importance of pediatric gonococcal (GC) infection in Florida and to describe its epidemiology, we reviewed surveillance data collected by the Florida Department of Health and Rehabilitative Services Sexually Transmitted Diseases program.

Methods

Public and private health care providers are required by law to report laboratory-confirmed episodes of GC infection to the Florida STD Program. The number of cases reported is computed each year by race (White and other-than-White) and sex, by age (0 to 9 years and by 5-year age groups over 9 years). For cases that occur among children younger than 13 years, a confidential morbidity card is completed by STD investigators with name, address, age, sex, race, date of diagnosis, site of GC infection and treatment, and forwarded to the STD Program. We reviewed morbidity cards that reported episodes of GC infection diagnosed between January 1, 1984, and December 31, 1988, in children aged 1 month through 12 years. A cluster of childhood GC infection was defined as two or more cases reported in children living at the same address and diagnosed within a 3-month interval. A recurrence was defined as two treated episodes of GC infection in a child separated by a time interval of at least 1 month.

Incidence rates of GC infection episodes per 100 000 children per year were calculated for the age group 1 month through 9 years by sex, race, and county of residence grouped as nonmetropolitan (no city with more than 50 000 residents) and metropolitan (at least one city with more than 50 000 residents). The 1986 mid-year Florida population under 10 years was used for rate calculation. We calculated the ratio of cases in children aged 1 month through 9 years to GC cases in males older than 14 years reported for the same time interval as an indicator of the proportion of male adults with GC infection who may have perpetrated child sexual abuse. Rates and ratios were standardized for race by the direct method.

Results

Eight hundred twenty-two cases of GC infection in children aged 1 month through 9 years were reported from 1984 to 1988. The review of available morbidity cards identified 854 episodes in 849 children aged 1 month through 12 years; 87% of these GC infection episodes were genital, 79% occurred in females, and 84% occurred in other-than-Whites (Table 1). Numbers of cases tended to be similar by age for males but not for females (Figure 1). Twenty clusters of pediatric GC infection were identified (16 clusters of two cases and 4 clusters of three cases). Five children had recurrent infections.

The overall incidence rate for the age group 1 month through 9 years was 11 per 100 000 per year. Incidence rates were higher for girls (18.2) than for boys (4.1) and for other-than-Whites (38.4) than for Whites (2.5). The race-adjusted incidence rate was higher for nonmetropolitan counties (15.6) than for metropolitan counties

At the time of this study, Jean-Claude A. Desenclos was with the Division of Field Services, Centers for Disease Control, and was assigned to the Florida Department of Health and Rehabilitative Services. David Garrity and Jack Wroten are with the Sexually Transmitted Diseases Program, Florida Department of Health and Rehabilitative Services, Tallahassee.

Requests for reprints should be sent to Jean-Claude A. Desenclos, MD, European Centre for the Epidemiological Monitoring of AIDS, Hôpital National de Saint-Maurice, 14 rue du Val d'Osne, 94410 Saint-Maurice, France.

This paper was submitted to the Journal November 27, 1990, and accepted with revisions June 10, 1991.

TABLE 1—Pediatric Gonococcal Infection Cases (in Children Aged 1 Month through 12 Years), by Site of Infection and Demographic Characteristics, Florida 1984–1988

Site of Infection[a]	Sex		Race[a]		Age, y						Total
	Male	Female	White	Other	<1	2–3	4–5	6–7	8–9	10–12	
Genitals[b]	147	595	87	655	17	155	169	130	100	171	742
	(83)	(88)	(66)	(91)	(61)	(83)	(88)	(87)	(90)	(92)	(87)
Mouth	8	42	29	21	4	12	11	9	4	10	50
	(5)	(6)	(22)	(3)	(14)	(6)	(6)	(6)	(4)	(5)	(6)
Anus	4	12	4	12	1	2	4	4	2	3	16
	(2)	(2)	(3)	(2)	(4)	(1)	(2)	(3)	(2)	(2)	(2)
Eyes	11	8	7	12	6	10	1	2	0	0	19
	(6)	(1)	(5)	(2)	(21)	(5)	(1)	(1)	(0)	(0)	(2)
Multiple sites	7	17	5	19	0	7	5	5	5	2	24
	(4)	(3)	(4)	(3)	(0)	(4)	(3)	(3)	(5)	(1)	(3)
Total	177	677	132	719	28	186	193	150	111	186	854
	(21)	(79)	(16)	(84)	(3)	(22)	(23)	(18)	(13)	(22)	(100)

Note. Numbers in parentheses are percentages of the total.
[a]No information was available on site of infection and race for 3 cases.
[b]Includes disseminated gonococcal infection (n = 4) and pelvic inflammatory disease (n = 13).

(8.1). The ratio of cases in the age group 1 month through 9 years to cases in males older than 14 years was 1.5 times greater for other-than-Whites (4.9:1000) than for Whites (3.3:1000), and was more than two times higher in nonmetropolitan counties (9.7:1000) than in metropolitan counties (4.2:1000) (Table 2).

Discussion

The Florida STD surveillance system identified a substantial number of pediatric GC infections that were very likely to have been transmitted through sexual abuse; it also detected clusters and recurrences. Previous reports have indicated that perhaps less than 6% of child molestations are ever reported.[1] Although nonsexual transmission may occur,[5] STDs in children are most often the result of sexual abuse.[3] Thus pediatric GC surveillance data may provide specific indicators to assess the importance of child sexual abuse in a community.[6–9]

Our data provide a description of the problem statewide, and may differ by demographic characteristics from some hospital based case-series data.[4,8,10–12] However, the clinical presentation is consistent with previous reports of pediatric GC infection case series,[4,7,8,10–13] and the age and sex distribution are also similar to that reported in two hospital-based studies of sexually abused children.[14,15]

However, the Florida pediatric GC infection episodes were not representative of the occurrence of child sexual abuse by race. Because the prevalence of GC infection is higher in Florida for other-than-

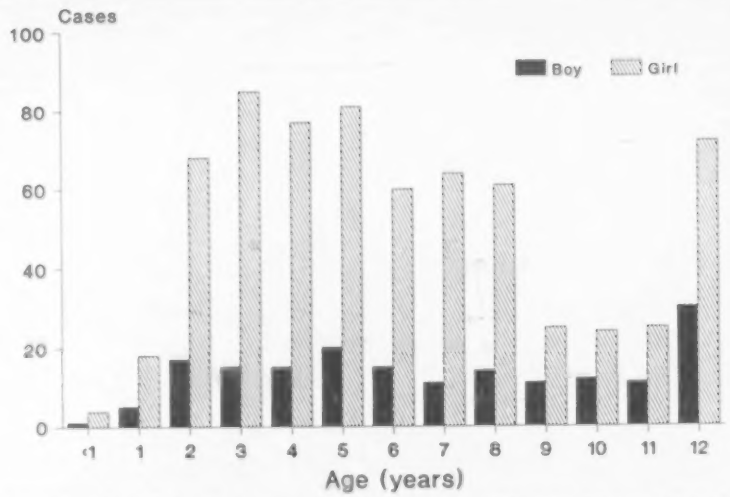


FIGURE 1—Cases of pediatric gonococcal infection (in children aged 1 month through 12 years), by age and sex, Florida 1984–1988.

TABLE 2—Ratio of Gonococcal Infection Cases in Children Aged 1 Month through 9 Years to 1000 Cases in Males Older Than 14 Years, by Race and County of Residence, Florida 1984–1988

Race	County of residence		
	Metropolitan	Nonmetropolitan	All counties
White	2.7	9.4	3.3
	(107:39 478)	(38:4050)	(145:43 528)
Other than Whites	4.6	9.8	4.9
	(586:127 231)	(94:9613)	(680:136 844)
Total[a]	4.2	9.7	4.6
	(693:166 709)	(132:13 663)	(825:180 372)

Note. Actual ratios are shown in parentheses.
[a] Standardized for race by the direct method.

Whites than for Whites,[16] the difference of pediatric gonorrhea incidence by race may not reflect a difference in the occurrence of child molestation. The child-to-adult case ratio, an attempt to adjust for the prevalence of GC infection in adult males (an indicator of the number of male adults with GC infection who may have perpetrated child sexual abuse), were mostly similar by race. Because it has not been used and validated previously, this ratio must be interpreted with caution; it assumes that all pediatric cases are sexually transmitted and that each perpetrator with gonorrhea infects only one child. However, the ratio was in the order of magnitude of the rates of confirmed child sexual abuse (of children younger than 15 years) investigated in Florida during fiscal year 1989 (Whites, 6121 victims, rate = 3.4 per 1000; non-Whites, 2043 victims, rate = 3.7 per 1000 [Florida Department of Health and Rehabilitative Services, Child, Youth and Family Services, unpublished data, 1989]). A bias toward a higher rate for other-than-Whites is possible, because other-than-Whites seek health care at public clinics more often than do whites[16]; therefore, their reportable diseases may be more likely than those of Whites to actually be reported. Pediatric GC rates were higher for nonmetropolitan areas than for metropolitan areas, a finding not previously reported.

Guidelines recommend that the identity of an STD agent isolated from a victim of sexual child abuse be confirmed by recognized laboratory methods.[2,3] Because they were not validated, the data presented in this paper need to be considered carefully. However, pediatric STD reporting systems may represent an alternative method of surveillance for child sexual abuse and warrant further evaluation. □

References

1. Fuller AK. Child molestation and pedophilia. *JAMA*. 1989;261:602–605.
2. Whittington BA, Rice RJ, Biddle JW, Knapp JS. Incorrect identification of *Neisseria gonorrhoeae* from infants and children. *Pediatr Infect Dis J*. 1988;7:3–10.
3. Centers for Disease Control. STD treatment guidelines. *MMWR*. 1989;38:5–8.
4. Beck-Sague C, Alexander ER. Sexually transmitted diseases in children and adolescents. *Infect Dis Clin N Am*. 1987;1:277–304.
5. Shore WB, Winkelstein JA. Nonvenereal transmission of gonococcal infections to children. *J Pediatr*. 1971;79:661–663.
6. Sgroi SM. "Kids with clap": Gonorrhea as an indicator of child sexual assault. *Victimology*. 1977;2:251–267.
7. McClure EM, Stack MR, Tanner T, Thevenin J, Gofstein RM, Helgerson SD. Pharyngeal culturing and reporting of pediatric gonorrhea in Connecticut. *Pediatrics*. 1986;78:509–510.
8. Branch G, Paxton P. A study of gonococcal infection among children and infants. *Public Health Rep*. 1965;80:347–352.
9. Sgroi SM. Pediatric gonorrhea and child sexual abuse: the veneral disease connection. *Sex Transm Dis*. 1982;9:154–156.
10. White TW, Loda FA, Ingram DL, Pearson A. Sexually transmitted diseases in sexually abused children. *Pediatrics*. 1983;72:16–20.
11. Groothuis JR, Bischoff MC, Jauregui LE. Pharyngeal gonorrhea in young children. *Pediatr Infect Dis J*. 1983;2:99–101.
12. Nair P, Glazer-Semmel E, Gould C, Ruff E. *Neisseria gonorrhoeae* in asymptomatic prepubertal household contacts of children with gonococcal infection. *Clin Pediatr*. 1986;25:160–163.
13. Folland DS, Burke RE, Hinman AR, Schaffner W. Gonorrhea in preadolescent children: an inquiry into source of infection and mode of transmission. *Pediatrics*. 1977;60:153–156.
14. Tilelli JA, Turek D, Jaffe AC. Sexual abuse of children: clinical findings and implications for management. *N Engl J Med*. 1980;302:319–323.
15. De Jong AR, Hervada AR, Emmet GA. Epidemiologic variation in childhood sexual abuse. *Child Abuse Negl*. 1983;7:155–162.
16. Florida Department of Health and Rehabilitative Services, Sexually Transmitted Disease Program. *Florida Sexually Transmitted Disease Statistics 1989*. Tallahassee, Fla: Florida Department of Health and Rehabilitative Services; 1990.

Evaluation of an Intervention to Reduce Playground Hazards in Atlanta Child-Care Centers

ABSTRACT

In 1988, we identified playground hazards at 58 child-care centers (CCCs) and intervened by showing the director the hazards and distributing safety information. In 1990, we evaluated the 58 intervention CCCs as well as 71 randomly selected control CCCs. Intervention centers had 9.4 hazards per playground; control centers had 8.0. We conclude that the intervention was ineffective. (*Am J Public Health.* 1992;82:429–431)

Jeffrey J. Sacks, MD, MPH, Mary D. Brantley, MPH, Patricia Holmgreen, MS, and Roger W. Rochat, MD

Introduction

A 1988 study of child-care centers (CCCs) found that the likelihood of a playground-related, medically attended injury increased as the number of identified playground hazards increased.[1] Part of that study involved an intervention in which we showed the director each identified hazard and distributed playground safety information. In 1990, we revisited these centers to see if the hazards had been corrected.

Methods

Methods for the sample selection and hazard survey have been described.[1,2] Briefly, 66 of 71 randomly selected CCCs enrolled in an injury study[2] underwent evaluation for commonly recognized playground hazards[3–6] (see also Table 1) between July and October 1988 (the baseline survey).[1]

The intervention consisted of pointing out and explaining each hazard to the CCC director, and distributing results of the injury study[2] and handbooks on playground safety.[7,8]

We had two questions: (1) would the intervention result in less hazardous playgrounds (intervention vs control study)? and (2) would specific hazards previously pointed out to CCC directors in 1988 be eliminated (persistence study)?

In April and May 1990, 16 CCC licensing inspectors with the Georgia Department of Human Resources Office of Regulatory Services (ORS) were trained in playground evaluation. Between May and June 1990, two-person teams of these inspectors evaluated all baseline survey CCCs and 71 randomly selected control CCCs (the follow-up survey). Inspectors were not told which CCCs were controls. Inspectors drew a map, inventoried equipment, measured equipment heights and undersurface thickness, listed hazards, noted the director's name and length of incumbency, and formed an overall impression of playground safety (very hazardous, hazardous, average, safe, very safe).

We also reviewed ORS annual licensing evaluations for 1988–1989 for all follow-up survey CCCs.

Intervention vs Control Study

We compared mean and median number of hazards per center and playground, and the inspector's impression of safety. Intervention CCCs were categorized by change in directorship after the intervention in 1988; control CCCs, by a change after September 1, 1988.

Persistence Study

Intervention center playgrounds and equipment were linked to the baseline survey by location and equipment height; those that could not be linked were excluded. Hazards were considered persistent if baseline and follow-up coding categories matched exactly.

Results

Intervention vs Control Study

Of 66 baseline survey CCCs, 58 were open in 1990. These CCCs had 123 playgrounds with 1171 pieces of equipment; the 71 control CCCs had 135 playgrounds and 1021 pieces of equipment. Climbing equipment was the most common piece of equipment. Of 238 climbers at least 4 feet tall, 125 (52.5%) were sited over nonresilient surfaces (earth or grass, concrete);

Jeffrey J. Sacks and Patricia Holmgreen are with the Division of Injury Control, National Center for Environmental Health and Injury Control, Centers for Disease Control in Atlanta; Mary D. Brantley and Roger W. Rochat are with the Office of Epidemiology at the Georgia Department of Human Resources in Atlanta.

Requests for reprints should be sent to Jeffrey J. Sacks, MD, MPH, Medical Epidemiologist, Division of Injury Control (F–36), Centers for Disease Control, 1600 Clifton Road, NE, Atlanta, GA 30333.

This paper was submitted to the journal April 9, 1991, and accepted with revisions August 22, 1991.

TABLE 1—Number and Mean Number of Playground Hazards[a] for 58 Intervention and 71 Control Child-Care Centers (CCCs), Atlanta, Ga, 1990

Type of Hazard[b]	Number (Mean) Hazards per Center		
	Intervention	Control	Total
Unrelated to equipment			
Off-limits access	49 (0.8)	63 (0.9)	112
Tripping hazard	54 (0.9)	46 (0.7)	100
Broken fence	40 (0.7)	45 (0.6)	85
Sharp fence wire	36 (0.6)	41 (0.6)	77
Exposed root	37 (0.6)	28 (0.4)	65
Exposed concrete	32 (0.6)	29 (0.4)	61
Exposed rocks	26 (0.5)	33 (0.5)	59
Debris	25 (0.4)	28 (0.4)	53
Broken/unsecurable gate	26 (0.5)	35 (0.5)	61
Briars	28 (0.5)	23 (0.3)	51
Nail	27 (0.5)	16 (0.2)	43
Splintered wood	23 (0.4)	19 (0.3)	42
Other	60 (1.0)	71 (1.0)	131
Total	463 (8.0)[c]	477 (6.7)[c]	940
Equipment-related			
Exposed concrete anchoring	71 (1.2)	67 (0.9)	138
Inadequate clearance	58 (1.0)	56 (0.8)	114
Rocks in fall zone	54 (0.9)	49 (0.7)	103
Exposed nail	47 (0.8)	43 (0.6)	90
Splintered wood	50 (0.9)	39 (0.6)	89
Open "S" hooks on swings	44 (0.8)	44 (0.6)	88
Rusting	27 (0.5)	28 (0.4)	55
Exposed root in fall zone	26 (0.5)	25 (0.4)	51
Sharp objects	29 (0.5)	25 (0.4)	54
Tip-over hazard[d]	25 (0.4)	24 (0.3)	49
Other	179 (3.1)	189 (2.7)	368
Total	610 (10.5)[e]	589 (8.3)[e]	1199

[a]Not including inadequate undersurfacing in fall zones.
[b]Data shown for those hazards recorded 40 or more times in the 129 centers.
[c]Standard deviation (SD) = 5.4 for intervention and 5.5 for control CCCs.
[d]Climbing equipment accounted for 10 (6 intervention, 4 control) of these hazards.
[e]SD = 9.2 for intervention and 8.7 for control CCCs.

only 17 (7.1%) had 3 or more inches of loose-fill undersurfacing.

Intervention and control CCCs had similar licensed capacities. Their playgrounds were designed for similarly aged children. The mean incumbency for intervention center directors was 8.4 years; for control center directors, 8.6 years. No control center directors had received the intervention.

The inspectors identified 2139 hazards (mean = 16.6, median = 13, range = 1–69) in the 129 CCCs. The median number of hazards per playground was 7.0; only 14 of 258 playgrounds had no hazards.

Intervention CCCs had 1.3 more hazards unrelated to equipment than control CCCs had and a mean of 2.2 more equipment-related hazards (Table 1). Intervention CCCs also had higher hazard scores (Table 2). The 32 centers with no change in directorship had higher hazard scores than the 26 centers with a change. Control CCCs with no change in directorship also had more hazards per playground than those with a change (mean = 9.3, median = 8 vs mean = 7.3, median = 6, respectively). As licensed capacity increased, playground hazard scores decreased (Table 2).

Of 118 intervention playgrounds, 47 (39.8%) were deemed "very safe" or "safe," 42 (35.6%) were "average," and 29 (24.6%) were "hazardous" or "very hazardous"; the corresponding figures for 127 control playgrounds were 43 (33.9%), 44 (34.6%), and 40 (31.5%). These differences were not statistically significant.[9] For intervention CCCs with no change in directorship, 28.8% of 59 playgrounds were rated "hazardous" or "very hazardous" compared with 20.3% for CCCs with a change. For control CCCs with no change in directorship, 41.5% of playgrounds were rated "hazardous" or "very hazardous" compared with 23.5% for CCCs with a change.

Licensing evaluation reviews revealed that, between 1988 and 1989, playground hazards increased from 85 to 104 for 53 intervention CCCs and decreased from 81 to 73 for 57 control CCCs. The 29 intervention CCCs with no change in directorship went from 49 hazards in 1988 to 63 in 1989. Of the 177 hazards noted in 1989, 41 (23.2%) had been cited in 1988.

Persistence Study

We linked 111 (90.2%) of 123 intervention playgrounds and 589 (51.7%) of 1140 pieces of equipment to baseline. At follow-up, we identified 150 (36.0%) of 417 hazards noted at baseline. For CCCs with a change in directorship, 48 of 153 hazards persisted (31.4%) compared with 102 of 264 hazards (38.6%) for CCCs with no change.

Discussion

At CCCs, most injuries requiring medical attention occur on playgrounds.[2,10–12] To reduce such injuries, we intervened, showing each director the specific hazards on the playground, explaining why they were problematic, and distributing educational materials.

Despite this intervention, quantitative scores indicate that intervention CCC playgrounds were not less hazardous than control playgrounds. Similar results were obtained from independent data—the annual licensing evaluations.

If the intervention had been effective, we would have expected CCCs with the same "intervention-exposed" director to have lower hazard ratings than CCCs with new, "unexposed" directors. The opposite occurred. Hazard persistence rates and subjective safety ratings were also worse in CCCs with intervention-exposed directors. This does not imply that the intervention had a negative effect. Rather, lower hazard ratings in control CCCs with changes in directorship suggest that leadership changes may be important in reducing hazards.

Almost 2 years after the intervention, 36% of the hazards found at baseline persisted. We cannot assume that the remaining 64% of the hazards were abated. The 1990 survey had 16 evaluators; the 1988 survey had only 3. What one team labeled an "exposed rock," another team may have called a "tripping hazard." Because we counted only exact matches, differences in nomenclature probably led to an undercount of true persistence.

Regulations do not ensure "hazard-free" environments. Although they are not defined, hazardous conditions are prohibited on center grounds. Nevertheless,

inspectors identified 2139 hazards in 129 CCCs. Georgia codes also mandate resilient surfaces beneath climbers without defining resilient. For climbers at least 4 feet tall, earth or grass is not considered a suitable undersurface,[13-14] and loose-fill undersurfaces (sand, wood chips, etc.) require 6 or more inches in depth.[3,13-16] Yet few CCCs followed these recommendations. As a result of this study, the State of Georgia has revised CCC regulations (effective March 1, 1991) to provide specific definitions of selected playground hazards.

The lack of better hazard scores in intervention CCCs, the better hazard ratings for intervention CCCs with intervention-*unexposed* directors, and the higher hazard persistence rates in CCCs with intervention-exposed directors support our conclusion that pointing out and explaining playground hazards to CCC directors, coupled with distributing educational materials, does not reduce hazards. Alternative approaches that need evaluation include rewriting regulations, training agency inspectors better, enforcing regulations more rigorously, educating CCC directors more extensively, and increasing parental awareness—for example, by posting inspection results in a conspicuous place in the CCC. □

TABLE 2—Hazard Scores of Intervention and Control Child-Care Centers (CCCs), by Change in Directorship, Atlanta, Ga, 1990

	Mean (Median) Scores				
	Intervention CCCs				
Number of Hazards	Same[a]	New[b]	Total	Control CCCs	All CCCs
Per CCC[c]	19.3 (15.5)	17.6 (13)	18.5 (15)	15.0 (11)	16.6 (13)
Per playground, by licensed capacity					
<60[d]	12.2 (12)	10.8 (10.8)	12.0 (11)	9.5 (8)	10.6 (9)
60–89[e]	6.2 (7.3)	8.3 (8)	7.3 (7.3)	9.5 (9)	8.3 (7.7)
90–119[f]	9.5 (8)	8.6 (7.8)	9.1 (8)	6.8 (6.5)	7.9 (7)
≥120[g]	13.5 (13.5)	7.9 (6.1)	9.0 (6.5)	6.5 (6.5)	7.3 (6.5)
Total[h]	10.1 (9)	8.5 (9)	9.4 (9)	8.0 (7)	8.6 (7.5)

[a]Director who received the educational intervention in 1988 is still the director in 1990.
[b]Director who received the educational intervention in 1988 is no longer the director in 1990.
[c]Standard deviations (SDs) = 14.3, 12.6, 13.5, 12.6, 13.1.
[d]SDs = 6.1, 1.8, 5.7, 5.5, 5.7.
[e]SDs = 3.6, 3.0, 3.4, 5.5, 4.6.
[f]SDs = 6.0, 5.8, 5.7, 5.1, 5.5.
[g]SDs = 13.4, 7.4, 8.2, 4.1, 5.8.
[h]SDs = 6.3, 5.3, 5.8, 5.2, 5.5.

Acknowledgments

Portions of the results of this study were presented at a National Institute of Child Health and Human Development workshop in Washington, DC, on March 25, 1991.

The authors thank the 16 Office of Regulatory Services (ORS) inspectors who surveyed the playgrounds; Marjorie Smith, RN, ORS, for providing resources, overall coordination, and impetus for the study; Jo Cato, Day Care Licensing Unit, ORS, for providing staff support and encouragement; Kenneth W. Holt, MSEH, Centers for Disease Control (CDC), and J. Marion Brown, Jr, BS, Georgia Dept of Human Resources (DHR), for field training of evaluators; J. David Smith, BS, and R. Keith Sikes, DVM, MPH, Office of Epidemiology, DHR for support and consultation; Peggy Monkus, BS, Office of Epidemiology, for coding data forms; Owen J. Devine, MS, CDC, for statistical consultation; and Suzanne M. Smith, MD, James A. Mercy, PhD, Richard W. Sattin, MD, and Suzanne C. Binder, MD, CDC, for editorial assistance.

References

1. Sacks JJ, Holt KW, Holmgreen P, Colwell LS, Jr, Brown JM, Jr. Playground hazards in Atlanta child care centers. *Am J Public Health.* 1990;80:986–988.
2. Sacks JJ, Smith JD, Kaplan KM, Lambert DA, Sattin RW, Sikes RK. The epidemiology of injuries in Atlanta day-care centers. *JAMA.* 1989;262:1641–1645.
3. US Consumer Product Safety Commission. *A Handbook for Public Playground Safety:* Vol 1. *General Guidelines for New and Existing Playgrounds.* Washington, DC: US Consumer Product Safety Commission; 1986.
4. Aronson SS, Aiken LS. Compliance of child care programs with health and safety standards: impact of program evaluation and advocate training. *Pediatrics.* 1980;65:318–325.
5. Hogan P. *The Playground Safety Checker: A Checklist Approach to Risk Management.* Phoenixville, Pa: Playground Press; 1988.
6. Werner P. Playground injuries and voluntary product standards for home and public playgrounds. *Pediatrics.* 1982;69:18–20.
7. US Consumer Product Safety Commission. *Play Happy, Play Safely: A Child-Centered Playground Safety Curriculum Approach.* Washington, DC: US Consumer Product Safety Commission; 1981.
8. US Consumer Product Safety Commission. *Play Happy, Play Safely: Playground Equipment Guide.* Washington, DC: US Consumer Product Safety Commission; 1979.
9. Fleiss JL. *Statistical Methods for Rates and Proportions.* New York, NY: John Wiley & Sons, Inc; 1981:150–157.
10. Landman PF, Landman GB. Accidental injuries in children in day care centers. *Am J Dis Child.* 1987;141:292–293.
11. Chang A, Lugg MM, Nebedum A. Injuries among preschool children enrolled in day care centers. *Pediatrics.* 1989;83:272–277.
12. Aronson SS. Injuries in child care. *Young Child.* 1983;38:19–20.
13. Reichelderfer TE, Overbach A, Greensher J. Unsafe playgrounds. *Pediatrics.* 1979;64:962–963.
14. Sweeney TB. X-rated playgrounds? *Pediatrics.* 1979;64:961.
15. US Consumer Product Safety Commission. *Playground Surfacing: Technical Information Guide.* Washington, DC: US Consumer Product Safety Commission; 1990.
16. Ramsey LF, Preston JD. *Impact Attenuation Performance of Playground Surfacing Materials.* Washington, DC: US Consumer Product Safety Commission; March 1990.

Air Quality during the Winter in Québec Day-Care Centers

Serge Daneault, MD, MSc, Monique Beausoleil, BSc, and Karen Messing, PhD

ABSTRACT

Over 90% of 91 day care centers in greater Montréal, Québec exceeded 1000 ppm of CO_2 during January through April 1989. Four variables were independent positive predictors of CO_2 levels: the density of children in the center; presence of electric heating; absence of a ventilation system; and building age. High levels of CO_2 are associated with respiratory tract and other symptoms. Clear standards and inspection policies should be established for day care center air quality. (*Am J Public Health.* 1992;82:432–434)

Introduction

Indoor air pollution in nonindustrial workplaces has recently become a focus of interest.[1–3] Studies of this phenomenon have concentrated on office buildings.[4–8] Recently, Scandinavian researchers[9–11] found inadequate ventilation in schools and day-care centers, especially those built during the 1980s.[9–11] Although small numbers of such centers were studied, these studies alert us to a problem potentially affecting the health of many very young children. In the United States, 18% of children under 6 were enrolled in day-care centers in 1982,[12] and this proportion is increasing.

In order to address this question, air quality was monitored in 91 day-care centers, representing 86% of such centers in the South Shore and eastern regions of Greater Montreal. The study was designed to assess air pollution and other physical conditions in the centers, and, in a second stage, to relate them to the health of the children attending the centers. We present here the assessment of physical conditions.

In this study, the temperature, percentage of relative humidity, and levels of carbon dioxide (CO_2) were chosen because they are (1) easy to measure, (2) reliable indicators of ventilation quality, (3) recommended by regulating agencies as the first step of tight building evaluations, and (4) used by Scandinavian researchers, providing a basis for comparison.[13–15] It should be noted, however, that toxic levels of unmeasured substances may persist in the presence of low levels of these indicators, particularly in the case of multiple-vocation buildings.

Methods

Choice of Day-Care Centers

A list of Montreal eastern region and South Shore day-care centers was derived from the lists of the provincial government registry that licenses such centers. In December 1988, a letter explaining the study was sent to each center, and this was followed by a phone call in January 1989 to request participation. Eighty-six percent participated in the study, almost all of which accept only children under 6.

General Characteristics of Centers

Each day-care center was visited on one afternoon between January 3 and April 15, 1989. Outdoor temperatures during this period ranged from −25°C to 16°C during the day, with a mean of 3°C. A technician noted, on a preformatted data sheet, characteristics of the building, the heating and air conditioning systems, ventilation and room division, and details of occupancy such as number of children per square meter of space and potential sources of air pollution.

Measurement of Indoor Air Characteristics

CO_2 was measured with an extra low range tube (Gastec Corporation) and relative humidity and temperature levels with a sling psychrometer (Bacharach Instruments). Measurements were taken in each of the five rooms most often occupied by children and were repeated four times: before, during, and after the afternoon nap and again before the children left. Because some day-care centers had fewer than five rooms, a total of 1672 measurements were taken. For 205 of these, CO_2 was measured simultaneously with an infrared gas analyzer (PM-2, ADC Company). Although the two measures

Serge Daneault is with the Center for the Study of Biological Interactions between Health and Environment, Université du Québec à Montréal, and Maisonneuve-Rosemont Hospital, Community Health Department, Montréal, Québec, Canada. Monique Beausoleil is with Maisonneuve-Rosemont Hospital, Community Health Department, Montréal, Québec, Canada. Karen Messing is with the Center for the Study of Biological Interactions between Health and Environment, Université du Québec à Montréal, Montréal, Québec, Canada.

Requests for reprints should be sent to Serge Daneault, MD, MSc, Charles Lemoyne Hospital, Community Health Department, 25 Taschereau, Bureau 100, Greenfield Park, Québec J4V 2G8 Canada.

This paper was submitted to the Journal September 7, 1990, and accepted with revisions May 14, 1991.

differed by an average of 32 ppm, in only four cases was there a difference greater than 300 ppm.

Data Analysis

The Statistical Package for Social Sciences (SPSS) program was used to calculate χ^2 values, and a stepwise multiple regression was constituted to predict carbon dioxide levels.

Results

CO_2 Levels

The CO_2 levels for the 91 day-care centers are shown in Table 1. There is no prescribed limit for day-care centers, but the average level for all 1672 measurements was 1505 ppm, well over the office building standard of 1000 ppm,[13] even considering the imprecision of the Gastech tubes. Only 9 of the 91 centers never exceeded 1000 ppm, and 13% exceeded 2500 ppm.

Relative Humidity and Temperature Levels

The relative humidity levels are reported in Table 2. The American Society for Heating and Refrigerating and Air Conditioning Engineers[16] recommends that humidity levels not be below 30% during winter. To our knowledge, there are not recommended levels specifically for day-care centers, but Canadian recommended winter levels for residences are 30% to 55%.[17] The levels of relative humidity were in the recommended range in 76 day-care centers (83%). Temperature was in the comfort zone (20°C to 23.6°C)[16] in 92% of the centers.

Predictors of CO_2 Levels

In order to focus efforts to diminish pollution, CO_2 levels were related to various building and occupancy characteristics. The results of the stepwise multiple regression are shown in Table 3. Four variables were significant and independent predictors of CO_2 levels. The main predictor was the density of children in the center when the measurements were made. Each child present per cubic meter increased the level of CO_2 by nearly 70 ppm. The use of an electric heating system (in 63% of the centers) increased the mean CO_2 level by 250 ppm. The presence of a ventilation system (in 32% of centers) reduced CO_2 levels by a mean of 562 ppm. Each year of building age decreased the level of CO_2 by 12 ppm. Thus, a building built in 1949 would be expected to have CO_2 levels 480 ppm lower than one built in 1989.

TABLE 1—Carbon Dioxide (CO_2) Levels in 91 Quebec Day-care Centers.

CO_2 level, [a]	Mean[b]		Minimum		Maximum	
	n	%	n	%	n	%
0–999	27	30	63	69	9	10
1000–2499	52	57	28	31	47	52
>2500	12	13	...	...	35	38
Total	91	100	91	100	91	100

[a]Average maximum CO_2 level = 2442 ppm; mean = 1505 ppm, and minimum = 861 ppm.
[b]Four measurements in each of five rooms per center.

TABLE 2—Relative Humidity in 91 Quebec Day-care Centers

Relative humidity, %[a]	Mean[b]		Minimum		Maximum	
	n	%	n	%	n	%
<30	...	...	7	8	...	...
30–39	21	23	43	47	3	3
40–54	55	60	38	42	40	44
≥55	15	17	3	3	48	53
Total	91	100	91	100	91	100

[a]Average maximum percentage of relative humidity = 55%, mean = 47%, and minimum = 39%.
[b]Four measurements in each of five rooms per center.

TABLE 3—Prediction of Mean Carbon Dioxide Levels by Building Characteristics of Day-care Centers

	Label	Step of Entry	β (SD)	R^2 Change
Density of children (per 100m^3)	Continuous	1	67.50 (19.91)	.1275
Heating system	1:Electric 0:Other	2	250.29 (181.12)	.0916
Artificial ventilation	0:No 1:Yes	3	−562.30 (179.14)	.0659
Age of building	Continuous	4	−12.07 (5.23)	.0512
Constant			1383.51	
Total R^2 = 0.34				

Note. The following additional variables were tested and not retained in the regression: building intended as day-care center (yes/no); renovation (yes/no); heating system controlled centrally (yes/no); hours of operation (coninuous); outdoor temperature, relative humidity, and wind speed (continuous).

Discussion

Air Quality in Day-Care Centers

Air quality was not good in the day-care centers. In a study in which Montreal day-care centers were very well represented, over 90% of centers exceeded 1000 ppm CO_2 at some time during the study, and 13% exceeded 2500 ppm. One sixth of the centers had humidity levels outside the accepted range. These results were obtained during afternoons of winter months in a northern region. That this may be a common problem is shown by the comparable results of Scandinavian researchers, who found levels over 4000 ppm in two day-care centers.[9]

Levels of CO_2 were related, as expected, to level of occupation, since CO_2 is produced mainly by human respiration. CO_2 level was also related to ventilation, either by a ventilating system or by natural air exchange in the case of older buildings. Since the variables had independent effects, we believe that the type of heating system and building age are not just sur-

TABLE 4—Energy Source for Heating as a Function of Building Age among 89 Day-care Centers in Quebec[a]

Energy Source	Age, y				
	<10	10–19	20–29	30–39	≥40
Electric	27	9	13	4	2
Oil	...	2	5	9	6
Gas	1	2	3	3	...
Other	2	...	...	1	...
Total	30	13	21	17	8

[a]This information was not available for two day-care centers.

rogate measures for the quality of building insulation. Electric heating is usually installed in new or recently renovated buildings (Table 4).

The level of CO_2 is important in itself, since it has been related to various symptoms, such as nasal discharge, eye irritation, throat dryness, headache, dry skin, and lethargy.[3,6,14] In addition, the level of CO_2 is an indicator of the quality of ventilation in general, and may thus serve as a marker for other pollutants, not verified in the present study, such as carbon monoxide and formaldehyde.[13,17]

Adequacy of Standards for Day-Care Centers

If the CO_2 levels and some of the humidity levels described in this study had been found in office buildings, immediate action would have been recommended according to the limits suggested by the American Society for Heating and Refrigerating and Air Conditioning Engineers.[13] Day-care center exposures not only involve workers, but also children. Also, the limit of 1000 ppm for CO_2 supposes an exposure time of 8 hours, while children and personnel may occupy day-care centers for as long as 12 hours per day.

Air quality standards appropriate to the use characteristics of day-care centers have not yet been developed. Buildings used for day care are only required to respect the current norms for buildings and occupancy generally in use in North America, in terms of numbers of children per unit area and height of ceilings. It may be that these standards are not adequate to protect the health of occupants of day-care centers. One study, for example, showed that children attending day-care centers suffered more illnesses than children at home.[18] Infectious diseases in crowded day-care centers may result from a synergistic effect of increased contact and poor ventilation, resulting in inadequate removal of microorganisms. If there is a link between these and other health effects and ventilation characteristics, criteria for day-care ventilation and building use should be adjusted. □

References

1. Wallace LA, Pellizzari ED, Hartwell TD, Davis V, Michael LC, Whitmore RW. The influence of personnel activities on exposure to volatile organic compounds. *Environ Res.* 1989;50:37–55.
2. Lowry S. Indoor air quality. *Br Med J.* 1989;299:1388–1390.
3. Samet JM, Marbury MC, Spengler JD. Respiratory effects of indoor pollution. *J Allergy Clin Immunol.* 1987;79:685–700.
4. Sterling TD, Collet CW, Sterling EM. Environmental tobacco smoke and indoor air quality in modern office work environments. *J Occup Med.* 1987;29:57–62.
5. Godish T. Indoor air pollution in offices and other non-residential buildings. *J Environ Health.* 1986;48:190–195.
6. Finnegan MJ, Pickering CAC, Burge PS. The sick building syndrome: prevalence studies. *Br Med J.* 1984;289:1573–1575.
7. Melius JK, Wallingford K, Keenlyside R, Carpenter J. Indoor air quality—the NIOSH experience. *Ann Am Conference Government Ind Hyg.* 1984;10:3–7.
8. Robertson AS, Burge PS, Hedge A, Sims J, Gill FS, Finnegan M. Comparison of health problems related to work and environmental measurements in two office buildings with different ventilation systems. *Br Med J.* 1985;291:373–376.
9. Lundquist GR, Iversen M, Dorsgaard J. Indoor climate in low-ventilated day-care institutions. *Environ Intern.* 1982;8:139–142.
10. Lundquist GR. Environmental studies in kindergartens. In: Proceedings of Healthy Buildings Symposia, September 5–8; 1988; Stockholm, Sweden.
11. Steensberg J. Indoor climate problems in day institutions for children. *Scand J Soc Med.* 1985;36(suppl):1–39.
12. Zigler E, Muenchow S. Infectious diseases in day care centers: parallels between psychologically and physically healthy care. *Rev Infect Dis.* 1986;8:514–520.
13. American Society for Heating and Refrigerating and Air Conditioning Engineers. Atlanta, GA. *Ventilation for Acceptable Indoor Air Quality.* American Society for Heating and Refrigerating and Air Conditioning Engineers; 1986.
14. Rajhans GS. Rationale for carbon dioxide guideline as an indicator of insufficient fresh air. In: Proceedings of the Fourth Canadian Building Congress, October 6–8, 1985; Ottawa, Canada.
15. Turiel I, Rudy J. *Occupant-generated CO_2 as an indicator of ventilation rate.* Washington, DC: US Department of Energy; 1980. Report LBL-10496, EEB-VENT 80-15.
16. American Society for Heating and Refrigerating and Air Conditioning Engineers. Atlanta, GA. *Thermal Environmental Conditions for Human Occupancy.* American Society for Heating and Refrigerating and Air Conditioning Engineers; 1981.
17. Ministère de la Santé Nationale et du Bien-être Social. *Directives d'exposition concernant la qualité des résidences. Rapport du comité consultatif fédéral-provincial sur l'environnement et l'hygiène du milieu.* Ottawa, Canada: Ministère de la Santé Nationale et du Bien-être Social; 1987:28–30.
18. Johansen AS, Leibowitz A, Waite LJ. Child care and children's illness. *Am J Public Health.* 1988;78:1175–1177.

Infant Mortality, Per Capita Income, and Adult Illiteracy: An Ecological Approach

R. Tresserras, J. Canela, J. Alvarez, J. Sentis, and L. Salleras

ABSTRACT

The ecological association of infant mortality rate (IM) with per capita income (PI) and prevalence of adult illiteracy (AI) has been studied using countries as the unit of analysis. A negative association between IM and PI in 1960 and 1982 has been observed (sample correlation coefficient [r] = –.625 and r = –.729, respectively; $P < .05$). A correlation between IM and AI has been found for both men and women ($r = .827$ and $r = .855$, respectively; $P < .05$). The ecological relative risks were 7.43 for men and 5.82 for women (95% confidence intervals: 5.16–10.71 and 4.36–7.75, respectively).

The association of IM and PI shows a slight but not significant improvement between 1960 and 1982. AI can be considered a good predictor of IM in countries. (*Am J Public Health.* 1992;82:435–438)

Introduction

Infant mortality (IM) has often been used as an indicator of community health level,[1] and it has been found to be associated with some socioeconomic and cultural variables.[2–10] Moreover, a significant correlation between IM, per capita income (PI); and adult literacy in 92 countries has recently been reported.[11]

The objectives of this paper are (1) to study the possible changes in the relationship between IM and PI, comparing data from 1960 and 1982; and (2) to assess the prevalence of adult illiteracy (AI) as an indicator of the ecological risk level of IM in a community.

Material and Methods

Data on IM[12] are from 1960 and 1982, data on PI are from 1962 and 1982, and data on AI are from 1980 and are calculated in real terms according to 1980 US dollar market prices.[13] This adjustment was made to avoid the effects of inflation and exchange changes.

For IM and AI, data were obtained from 103 countries; for PI, data were obtained from 82 countries for 1962 and from 95 countries for 1982. For each variable, all available data were included in each case. The 82 countries with available data on PI in 1962 were also included in the 95 countries with available data on PI in 1982, and all of these were included in the 103 analyzed countries.

To study how the association of IM and PI differed between 1960 and 1982, the difference in regression coefficients for both years was analyzed.[14] Correlation between observed and expected IM rates for 1982 was also studied, applying the 1960 regression equation.

To estimate the ecological relative risk (RR), the weighted logistic regression method proposed by Silcocks and Murphy[15] was used. The weight of each country was made according to the number of births, estimated from the total population and its crude birth rate. The weight was used to adjust for the relative influence of each country in the model, thereby preventing small or large countries from dominating the analysis.

Normal distribution of all variables was checked using the Kolmogorov-Smirnov test. Normalization of PI for the studied years was obtained by applying natural logarithms to original data.[16]

Results

Table 1 shows the main parameters for the studied variables. The correlation between IM and PI was statistically significant for 1960 and 1982 (sample correlation coefficient [r] = –.625 and r = –.729, respectively, $P < .05$) (Figures 1a and 1b). The regression equations were

$$1960: \text{IM} = 337.144 - 32.460 \cdot \ln \text{PI}$$

$$(R^2 = .3906) \quad (n = 82),$$

$$1982: \text{IM} = 275.543 - 28.234 \cdot \ln \text{PI}$$

$$(R^2 = .5315) \quad (n = 95).$$

Although PI may explain a greater IM variability for 1982 than for 1960, no differences in regression coefficients were found ($t = .83$). The correlation between observed and expected IM rates for 1982 was $r = .73$ ($P < .05$).

AI prevalence was greater in women than in men (43.3% vs 25.5%; $z = 3.85$; $P < .05$). The correlation between IM and AI was positive and statistically significant for both men and women ($r = .827$ and

R. Tresserras, J. Canela, J. Alvarez, and L. Salleras are with the Health Department, Public Health Services, Generalitat de Catalunya, in Barcelona; J. Sentis is with the Health Sciences Division, Bioestatistics Unit, University of Barcelona.

Requests for reprints should be sent to Ricard Tresserras i Gaju, Departament de Sanitat i Seguretat Social, Travessera de les Corts 131–159, 08028 Barceloua, Spain.

This paper was submitted to the Journal May 29, 1990, and accepted with revisions July 18, 1991.

TABLE 1—Infant Mortality (IM) Rates (1960 and 1982), Per Capita Incomes (PI) (1962 and 1982), and Prevalences of Adult Illiteracy (AI) (in Men and Women, 1980).

Variable	Mean/Median/ Weighted Prevalence	Minimum value	Maximum value	n
IM 1960 10^3	116[a]	26	230	103
IM 1982 10^3	84[a]	7	200	103
PI 1962 ($)	665[b]	117	21 131	82
PI 1982 ($)	1074[b]	74	21 314	95
AI men (%) (1980)	25.5[c]	1	86	103
AI women (%) (1980)	43.3[c]	1	98	103

Note.—n = number of countries with available data.
[a]Mean.
[b]Median.
[c]Weighted prevalence based on 1980 population.

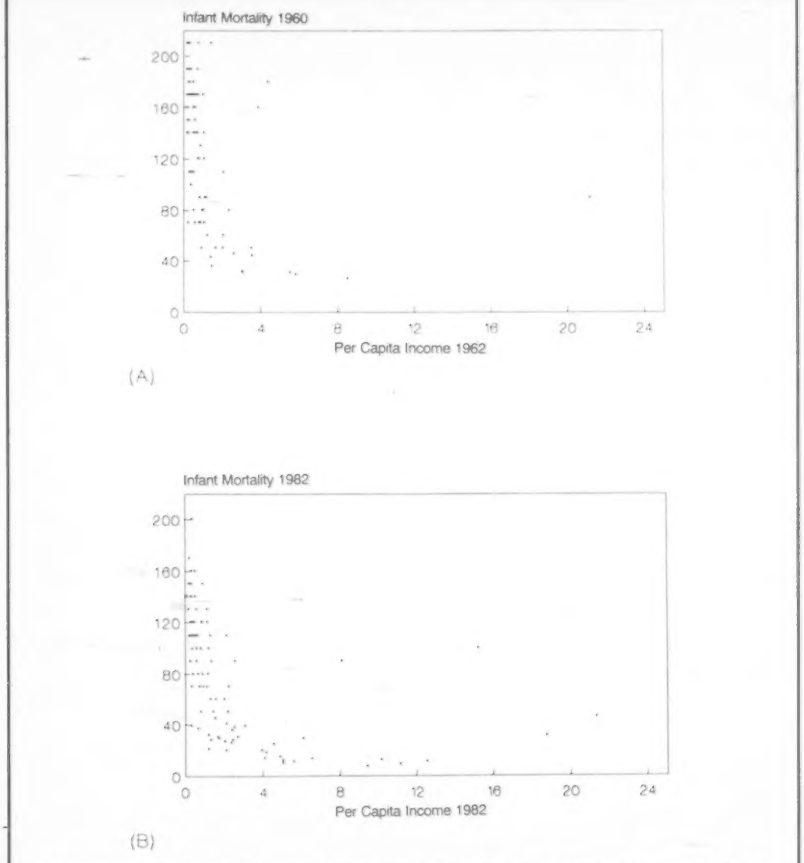


FIGURE 1—Correlation between infant mortality rate and per capita income (US dollars) in 1960 (1a) and 1982 (1b). PI in arithmetic scale.[12,13]

$r = .855$, respectively; $P < .05$) (Figures 2a and 2b). The regression equations were

$$\text{Men: IM} = 27.835 + 1.649 \cdot \text{AI} \quad (R^2 = .684),$$

$$\text{Women: IM} = 18.376 + 1.356 \cdot \text{AI} \quad (R^2 = .731).$$

The ecological relative risk was 7.43 for men and 5.82 for women (95% confidence intervals: 5.16–10.71 and 4.36–7.75, respectively).

Discussion

Although data could be obtained for only 103 countries, these countries represent approximately 90% of the world's population (4500 million people) (Figure 3).

The results support an association at the ecological level between IM and independent variables, such as PI and AI. Although there was an important reduction in IM rates from 1960 to 1982, IM is still considered high, being $84 \cdot 10^3$ in 1982.

Regarding the association of IM and PI, the observed 1982 IM rate was very similar to the expected rate, and there are no differences between regression coefficients although data show an improvement in absolute terms. This suggests that the association between these variables did not change significantly during the study period. However, a negative exponential association was observed. Figures 1a and 1b (in arithmetic scale) show that there is a vertical branch formed by developing countries and a horizontal one formed by industrialized countries, with the oil-exporting countries between them. It should also be noted that while maximum income increases, minimum income decreases.

A significant difference related to AI was observed between sexes. Although this type of difference is difficult to analyze in this kind of study, sex differences do not seem to explain the higher or lower IM risk in the studied countries. However, at the ecological level, countries with higher prevalence of AI among both men and women also show a significantly higher rate of IM. Furthermore, there is also an association between PI and AI, and the relationship between IM and AI is at least as strong as that existing between IM and PI. In a previous report,[11] we found that both PI and AI were included in a regression stepwise model.

Over the study period, developed countries showed the highest improvement in their IM rates. Only oil-exporting countries showed a similar proportional reduction, but they were still far behind despite their very high PI levels.

In interpreting the results, the possible presence of some biases has to be taken into account. These biases can be due to the countries selected (in this case, available data for all of them came from

the same source) or to the fact that data are estimated by international organizations. Moreover, PI does not reflect the wealth distribution of a country, nor does AI reflect the general cultural level of the different social groups. Finally, to avoid the ecologic fallacy,[17,18] these results cannot be regarded at an individual level.

If the above-mentioned aspects are considered, it can be said that PI and AI are variables that have to be taken into account in studying IM rates in countries. Relevant changes have not been found in the association of IM with PI between 1960 and 1982. AI can be considered a good predictor of IM in countries. □

Acknowledgments

The authors thank Prof. V. Navarro from The Johns Hopkins University, who encouraged the development of this paper; Prof. J. L. Tabcmer, Prof. Ll. Serra-Majem, C. Castell, E. Salto, and E. Sánchez from the Catalonian Health Department for their advice; and C. Garsot for the English review.

References

1. Lindsay E. The epidemiology of perinatal mortality. *World Health Stat. Q.* 1985;38:289–301.
2. Woolhandler S, Himmelstein DU. Militarism and mortality: an international analysis of arms spending and infant death rates. *Lancet.* 1985;1:1375–1378.
3. Robinson MJ, Palmer SR, Avery A, James CE, Beynon JL, Taylor RW. Ethnic differences in perinatal mortality: a challenge. *J Epidemiol Community Health.* 1982;36:22–26.
4. Wise PH, Kotelchuck M, Wilson ML, Mills M. Racial and socioeconomic disparities in childhood mortality in Boston. *N Engl J Med.* 1985;313:360–366.
5. Stockwell EG, Swanson DA, Wicks JW. Economic status differences in infant mortality by cause of death. *Public Health Rep.* 1988;103:135–142.
6. Elbourne D, Pritchard C, Dauncey M. Perinatal outcomes and related factors: social class differences within and between geographical areas. *J Epidemiol Community Health.* 1986;40:301–308.
7. Barell V, Wax Y, Ruder A. Analysis of geographic differentials in infant mortality rates: the Or Yehuda community. *Isr J Epidemiol.* 1988;128:218–230.
8. Bradshaw J, Edwards H, Lawton D, Staden F, Weale J, Weekes A. Area variations in infant mortality 1975–77. *J Epidemiol Community Health.* 1982;36:11–16.
9. Parazzini F, La Vecchia C. Perinatal and infant mortality rates and place of birth in Italy, 1980. *Am J Public Health.* 1988;78:706–707.
10. Salleras L, Sentis J, Canela J, Garcia A. Facteurs sociaux et d'assistance sanitaire et mortalité périnatale en Espagne: etude écologique. *Rev. Epidemiol Santé Publique.* 1988;36:30–35.
11. Tresserras R, Alvarez J, Canela J, Sentis J, Salleras L. Mortalidad infantil, renta per cápita y alfabetización: asociación ecológica a nivel internacional. *Cuad Bioestat Apli Infor.* 1990;8:32–43.
12. Grant JP. *The State of the World's Children 1985.* New York, NY: Oxford University Press; 1985. Published for the UN Children's Foundation.
13. International Monetary Foundation. *International Finantial Statistics: Supplement on Output Statistics.* Washington, DC: International Monetary Foundation. 1984;18–21. Supplement Series 8.
14. Armitage P, Berry G. *Statistical Methods in Medical Research.* 2nd ed. Oxford, England: Blackwell Scientific Publications; 1987:273–282.

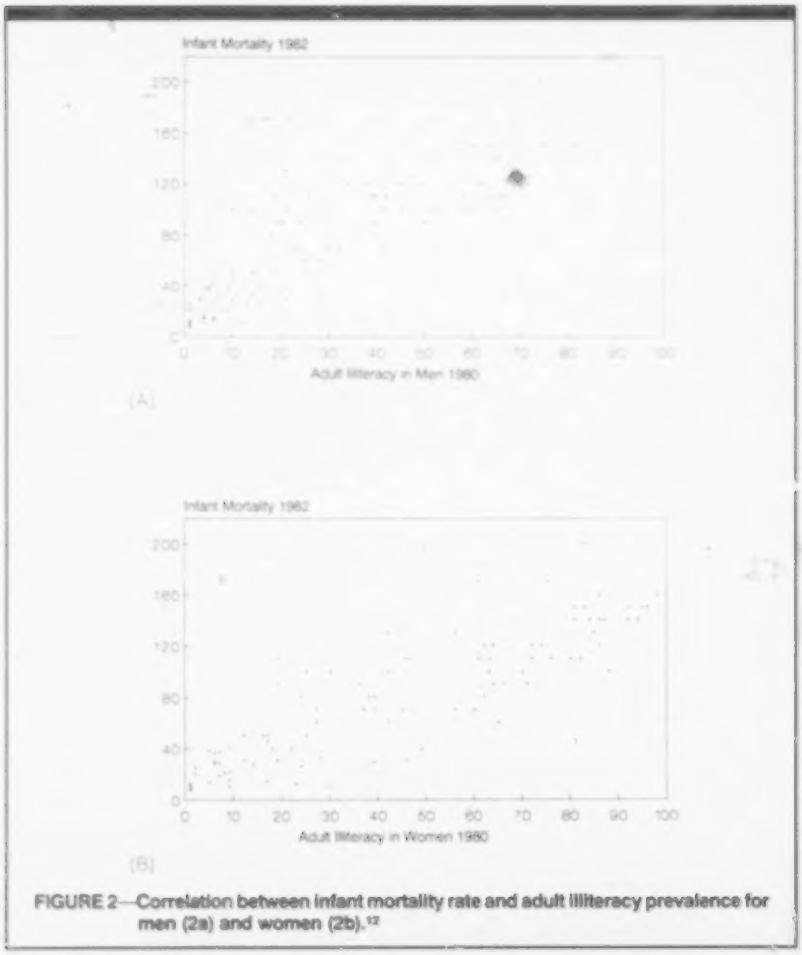


FIGURE 2—Correlation between infant mortality rate and adult illiteracy prevalence for men (2a) and women (2b).[12]

FIGURE 3—Countries marked in black are those for which data were available.

15. Silcocks PBS, Murphy M. Relative risk estimation from vital statistical data: validation, a pitfall and an alternative method. *J Epidemiol Community Health.* 1987;41:59–62.
16. Schwartz D. *Méthodes statistiques à l'usage des médecins et des biologistes.* Paris, France: Flammarion; 1963:225–231.
17. Morgenstern H. Uses of ecologic analysis in epidemiologic research. *Am J Public Health.* 1982;72:1336–1344.
18. Piantadosi S, Byar DP, Green SB. The ecological fallacy. *Am J Epidemiol.* 1988; 127:893–904.

Which Newborns in New York City Are at Risk for Special Education Placement?

Doris Goldberg, MD, MPH, Margaret McLaughlin, MD, MPH, Margaret Grossi, MD, MPH, Alex Tytun, PhD, and Steve Blum, PhD

ABSTRACT

In this study of 162 third graders in New York City public schools, we found that slightly over half of the children in special education were males who had Medicaid coverage at birth and mothers with medical conditions or adverse health habits noted on the birth certificate; two thirds of the children with this combination of characteristics actually were placed in special education. These findings suggest that newborns at risk for later learning disabilities can be targeted to receive preventive interventions. (*Am J Public Health.* 1992;82:438–440)

Introduction

Public Law 99-457 supports a national initiative to establish early identification and service systems for children with special needs from birth to 3 years of age. Some of the children who will require special education services in the early grades of school can be detected as having special needs in the first 3 years of life. Some may have medical conditions associated with cognitive difficulties, and others may have health or social risks for later learning problems. In this pilot study, we linked the educational placement records of third graders in the New York City public school system with information on their birth certificates and retrospectively analyzed the association of special education placement with characteristics recorded on birth certificates. The identification of newborns at risk can supplement the current identification of children with special needs and allow children to be targeted for preventive interventions in the preschool years to reduce their need for costly special education services in school.

A combination of reproductive health and sociodemographic variables was found to be the strongest predictor of school failure when Ramey et al.[1] studied birth certificate data from North Carolina. Escalona[2] found increased jeopardy for socially disadvantaged children with biological risks. Chamberlin[3] recently concluded that perinatal factors have a decreasing impact and environmental factors have an increasing impact on developmental status as children progress from birth to the age of 7. Palfrey et al.[4] found an association between mother's educational attainment and early diagnosis of high- and low-prevalence learning disorders in children.

Methods

Data for this study consisted of linked birth and school records of 162 third graders attending New York City's public schools. Access to school health records was obtained in three elementary schools located in three demographically distinct school districts: Bedford-Stuyvesant, Coney Island, and Staten Island. All third graders in special education classes (n = 45) were treated as "cases"; more than twice this number (n = 117) were randomly selected from the general education third-grade populations in the same schools and treated as "controls." The case group was broadly defined to include the mildly retarded, the emotionally disturbed, and those requiring speech or other therapies to foster learning. Third graders were chosen because placement in special education classes is generally arranged at the end of second grade for children who exhibit learning or behavioral difficulties in the general classroom setting.

The dependent variable in this study was educational placement status; the predictor variables were obtained primarily from birth certificates, although additional information was gathered from available

The authors are with the New York City Department of Health.

Requests for reprints should be sent to Dr. Doris Goldberg, New York City Department of Health, 125 Worth Street, New York, NY 10013.

This paper was submitted to the Journal August 20, 1990, and accepted with revisions June 10, 1991.

school records. For example, "medical complications" is a constructed variable representing maternal pregnancy complications, chronic medical conditions preceding pregnancy, adverse health habits (e.g., smoking, alcohol or substance abuse), or a combination of these factors. "Changed school district" is a variable derived by comparing the community school district in which the residence address stated on the birth certificate is located with that where the child is actually enrolled in school; it is a rough indicator of the within-city mobility of children with special needs.

Results

Table 1 shows the percentage distribution of special- and general-education children across 12 variables; the distribution of the total number of live births in New York City during 1980 (the year of

TABLE 1—Percentage Distributions of the Study Sample and 1980 New York City (NYC) Live Births across Selected Demographic and Health Status Variables

	Study Sample		1980 NYC Live Births[a] (n = 107 000)
	Special Ed. (n = 45)	General Ed. (n = 117)	
Sex (male)	76	44	51
Race (Black)	60	38	35
Single mother	64	33	35
Medicaid coverage	71	34	32
Mother's education			
Elementary school	13	8	28
High school	76	66	38
Post–high school	9	26	22
Mother's age (>25 y)	57	47	39
Total deliveries (>4)	24	14	39
Prenatal care (late/none)	24	11	18
Medical complications (one or more)	29	11	7
Gestation (<37 wk)	19	8	11
Birth weight (<2500 g)	13	8	9
Changed school district	44	56	NA

Note. All percentages are rounded and are based on subjects for whom information is available on the variables considered. NA = not available.
[a]Source: New York City Department of Health, Bureau of Health Statistics and Analysis.

TABLE 2—Odds Ratios (ORs) for Selected Variables, by Educational Placement (Special or General) and Race: Stratum-Specific and Pooled Odds Ratios with 95% Confidence Intervals (CIs)

	Race					
	Black		Other		Homogeneity χ^2 (df = 1)	Pooled OR[a] with 95% CI
	Special	General	Special	General		
Sex[b]						
Male	22	16	12	36	2.7, $P > .05$	1.8 ≤ 3.9 ≤ 8.1
Female	5	28	6	37		
	(OR = 7.7)		(OR = 2.1)			
Single mother[b]						
Yes	20	31	9	8	5.8, $P < .02$	—
No	7	13	9	65		
	(OR = 1.2)		(OR = 8.1)			
Medicaid coverage[c]						
Medicaid	21	31	11	9	6.8, $P < .01$	—
Self/other	6	12	7	64		
	(OR = 1.4)		(OR = 11.2)			
Mother's education						
<12 y	24	31	16	52	0.0, $P < .90$	1.2 ≤ 3.4 ≤ 9.9
≥12 y	2	12	2	21		
	(OR = 3.5)		(OR = 3.2)			
Prenatal care[d]						
Late/none	8	9	3	4	0.6, $P < .40$	0.9 ≤ 2.1 ≤ 5.1
Early	18	34	14	67		
	(OR = 1.7)		(OR = 3.6)			
Medical complications[b]						
One or more	10	7	3	6	0.1, $P < .70$	1.2 ≤ 2.8 ≤ 6.7
None	17	37	15	67		
	(OR = 1.7)		(OR = 2.2)			
Gestation[d]						
<37 wk	5	7	3	2	2.9, $P > .05$	0.8 ≤ 2.0 ≤ 5.3
≥37 wk	22	37	13	68		
	(OR = 1.2)		(OR = 7.8)			
Birth weight[b]						
<2500 g	5	6	1	3	0.0, $P < .99$	0.5 ≤ 1.4 ≤ 4.4
≥2500 g	22	38	17	70		
	(OR = 1.4)		(OR = 1.4)			

[a]The pooled OR, the Mantel-Haenszel estimate, is not reported where the descriptive alpha level of the χ^2 test of homogeneity is less than .05. [b]n = 162. [c]n = 161. [d]n = 157.

TABLE 3—Exploratory Results of Stepwise Regression Analysis Identifying a Set of Minimally Redundant Variables That Can Predict Special Education Status

Iteratively reweighted least-squares (IRWLS)
Multiple linear logistic regression

Analysis of Variance

	df	Weighted Sum of Squares (χ^2)	Weighted Mean Square	P Value
Model (omnibus)	3	27.74	9.25	<.001
Residual (goodness-of-fit)	157	154.16	0.98	>.550
Total (corrected)	160	181.90		
Weighted intercept	1	12.91		
Uncorrected total	161[a]	194.80		

R^2 : .15249
Adjusted R^2 : .13630
Related chi-squared statistics

	χ^2	df	P Value
Score (omnibus)	35.75	3	<.001
Log likelihood (omnibus)	39.28	3	<.001
Log likelihood (goodness-of-fit)	151.50	157	>.600

Variables in the Equation

Risk Factor	b	SE_b	Partial χ^2	Unadjusted P value	Tolerance
MEDICAID	1.5540	.4161	13.95	<.001	.39516
SEX	1.7786	.4671	14.50	<.001	.27299
MEDCOMP	1.4403	.5382	7.16	<.010	.69377
Intercept	−3.1356				.15594

Note. MEDICAID: 1 = medicaid recipient at time of birth, 0 = otherwise; SEX: 1 = male, 0 = female; MEDCOMP: 1 = one or more medical complications at time of birth, 0 = none.
[a]Information was unavailable for one subject.

birth for the majority of the study sample) across the same variables is shown for comparative purposes. Special-education children were far more likely to be Black males whose unmarried mothers lacked formal education beyond high school and who had Medicaid coverage at the time they gave birth.

Presented in Table 2 are some of the same variables stratified by both race and educational placement. "Sex," "mother's education," and "medical complications" continue to show a strong association with educational placement (i.e., the pooled odds ratios exceed 2) when the effects of race are controlled statistically. However, the degree of association between educational placement and Single Mother or Medicaid Coverage is so different between racial groups as to dictate against estimating a common odds ratio. Among Black children in this study, Medicaid coverage appears to be unrelated to educational placement.

Stepwise linear logistic regression (Table 3) was performed with 12 key variables to identify a set of minimally redundant variables that could best predict educational placement. Regardless of whether forward selection or backward elimination methods were employed, or whether variables were treated as categoric or numeric, the same three factors emerged: sex of child (SEX), Medicaid coverage at time of birth (MEDICAID), and one or more medical complications at birth (MEDCOMP). Because this function was estimated solely for purposes of prediction, no substantive interpretation will be given to the regression coefficients.

Classifying children by their sex, their mother's Medicaid eligibility, and the presence or absence of medical complications at birth enabled us to correctly identify 25 out of 45 children in special education and, less importantly, 99 out of 116 children in general education.

Discussion

This pilot study suggests that certain health factors and sociodemographic characteristics that can be identified at birth are strongly associated with third graders' educational placement. It would be worthwhile to apply a similar kind of analysis to a much larger sample that would be more representative of the city's 32 community school districts. Specific risk factors can be expected to vary within diverse pupil populations, as was demonstrated in the present study with racial stratification.

Sex, Medicaid coverage, and medical complications, taken together, were the risk factors most strongly associated with educational placement in the regression analysis. Although controversy remains as to whether biological or sociological factors determine why more boys than girls end up in special education classes,[5] extensive data suggest that males are more likely to be candidates for special education for biological reasons.[6]

The preliminary results presented in Table 4 imply that one in every three children whom we can identify as being at risk for special education will eventually require it. This implication, in turn, suggests that prognostic methodologies such as the one described here can be employed to supplement the traditional indicators for medical assessment, rather than to replace accepted methods for screening and evaluating newborns and preschoolers with known biological risks. Classification procedures like those employed here can nevertheless aid in more precisely identifying newborns who should be followed continuously and screened periodically and who should receive supportive services. Such interventions should significantly reduce the chances that early vulnerability will progress to later disability, which will be more difficult and costly to treat. □

Acknowledgments

The authors thank Jean Lee, Marge Manley, and the cooperating school health staffs who made this study possible.

This paper was presented at Prevention '90 Conference in Atlanta, Ga, March, 1990.

References

1. Ramey CT, Stedman DJ, Borders-Patterson A, Mengel, W. Predicting school failure from information available at birth. *Am J Ment Defic*. 1978;6:525–534.
2. Escalona SK. Babies at double hazard: early development of infants at biologic and social risk. *Pediatrics*. 1982;70:670–676.
3. Chamberlin RW. Developmental assessment and early intervention program for young children: lessons learned from longitudinal research. *Pediatr Rev*. 1987;8:237–247.
4. Palfrey JS, Singer J, Walker D, Butler J. Early identification of children's special needs: a study of five metropolitan communities. *J Pediatr*. 1987;111:651–659.
5. Shaywitz S, Towle V, Kesse D, Shaywitz B. Prevalence of dyslexia in boys and girls in an epidemiologic sample. *JAMA*. 1990;181:143–157.
6. Funnucci J, Childs B. Are there really more dyslexic boys than girls? In: Ansara A, ed. *Sex Differences in Dyslexia*. Baltimore, Md: The Orton Society; 1981.

The Lowest Birth-Weight Infants and the US Infant Mortality Rate: NCHS 1983 Linked Birth/Infant Death Data

Mary D. Overpeck, MPH, Howard J. Hoffman, MA, and Kate Prager, ScD

ABSTRACT

The National Center for Health Statistics Linked Birth and Infant Death Data Set, 1983 birth cohort, shows that infants weighing less than 750 g, comprising only 0.3% of all births, account for 25% of deaths in the first year of life and for 41% of deaths in the first week. If interventions had prevented the death of these very small babies, the infant mortality rate would have been 8.3 per 1000 live births instead of 10.9, and the Black/White mortality differential would have been reduced by 25%. (*Am J Public Health*. 1992;82:441-444)

Introduction

Low- and very low-birth-weight perinatal mortality has decreased dramatically over the past few decades.[1,2] California linked birth and death certificates show that the decline in perinatal mortality rates in the 1970s occurred primarily from improvements in the survival of infants born at weights above 750 g.[3] Virtually no change was observed for infants weighing under 751 g, and only a small decrease was observed for infants weighing between 751 and 1000 g. Prospective studies of reproductive outcomes also show higher losses among fetuses from 16 to 24 gestational weeks than among fetuses from 25 to 29 gestational weeks, with clinical assessment of very poor viability for fetuses under 500 g or 22 weeks of fertilization age.[4-8]

The Linked Birth and Infant Death Data Set, 1983 birth cohort, from the National Center for Health Statistics (NCHS) is the first national linkage of individual birth and death certificates since 1960.[9,10] Linkage overcomes limitations of infant mortality rates (IMRs), in which the numerator is based only on annual counts of death certificates; it also permits calculation of mortality rates by birth weight for the whole United States. A 1980 analysis based on linkages by individual states shows the impact of very low-birth-weight (VLBW) infants on the US IMR. This tabular data set, however, did not include births under 500 g or allow comparisons of infants born at weight intervals smaller than 500 g.[11]

The present analysis focuses on individually linked data for the least viable infants (those weighing less than 500 g and those weighing 500 to 749 g), their contribution to the total IMR, and the rate and age at which they die. Black and White births are analyzed separately because of differential rates of extremely low-weight births and attributable proportions of deaths.

Methods

The methodology to create the 1983 linked file is described in both the NCHS Public Use Data Tape Documentation and a working paper.[12,13] The file contains linked records for *all* infants born in 1983 in the United States who died in 1983, or in 1984 before their first birthday.

Records were linked for 39 683 resident infant deaths, 98.4% of all infant deaths occurring in the United States during the time specified. Analysis of the 651 unlinked death certificates shows them to be similar to the linked certificates in terms of race and age at death. If no death certificate was filed, an infant with a birth certificate was assumed to have survived. Specific rates may be slightly understated owing to unlinked records or to missing information on the linked records. Birth weight is missing for 2.9% of linked infant deaths and for 0.1% of all births. Birth weight is used here instead of gestational age because it is more completely and accurately reported.[14,15]

Differences larger than 2% in Figures 1 and 2 are statistically highly significant based on Fleiss' confidence interval procedures.[16] Otherwise, statistical significance has been assessed using chi-square tests on relative risk (RR) estimates, as discussed for Table 1.

Results

Forty-three percent of all infant deaths came from 1.2% of all live births, the VLBW infants weighing less than 1500 g (Table 1). The majority of deaths among VLBW infants came from those weighing 500 to 749 g and those weighing less than 500 g, with 25% of total deaths coming

Mary Overpeck and Howard Hoffman are with the National Institute of Child Health and Human Development, National Institutes of Health, in Bethesda. Kate Prager is with the National Center for Health Statistics, Centers for Disease Control, in Hyattsville.

Requests for reprints should be sent to Mary D. Overpeck, MPH, Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Human Development, Executive Plaza North, Room 640, Bethesda, MD 20892.

This paper was submitted to the Journal February 19, 1991, and accepted with revisions July 18, 1991.

TABLE 1—Number and Percent of Total US Live Births and Infant Deaths and Mortality Rate per 1000 Births, by Race and Birth Weight: 1983 Birth Cohort

	Live Births				Infant Deaths						
		Percent			Percent			Rate/1000 Live Births			Relative Risk:
	Total Number[a]	Total[a]	White	Black	Total[a]	White	Black	Total[a]	White	Black	Black/White
Total	3 639 113	100.0	79.8	16.1	100.0	68.3	27.9	10.9	9.3	18.9	2.03
2500 g or more	3 385 912	93.0	75.2	14.1	38.7	28.1	8.8	4.5	4.1	6.8	1.66
1500–2499 g	204 534	5.6	3.8	1.6	15.4	10.9	3.9	30.0	31.4	26.6	0.85
<1500 g	43 284	1.2	0.7	0.4	43.0	27.4	14.2	393.6	403.4	378.8	0.94
750–1499 g	31 246	0.9	0.5	0.3	18.4	12.7	5.0	233.1	253.3	194.5	0.77
500–749 g	7594	0.2	0.1	0.1	14.7	8.9	5.3	765.7	790.1	735.6	0.93
<500 g	4444	0.1	0.1	0.1	9.9	5.8	3.9	885.9	890.6	877.1	0.98
Weight not stated[b]	5383	0.1	0.1	0.0	2.9	1.8	1.0	216.1	176.5	381.9	2.16

Note.—Percents may not sum to 100 due to rounding.
Source.—National Center for Health Statistics, National Vital Statistics System (unpublished data).
[a]Includes race other than White and Black and "not stated" weights; total number of deaths is shown in Table 2.
[b]Deaths for which birth weight was not stated on the linked birth certificates.

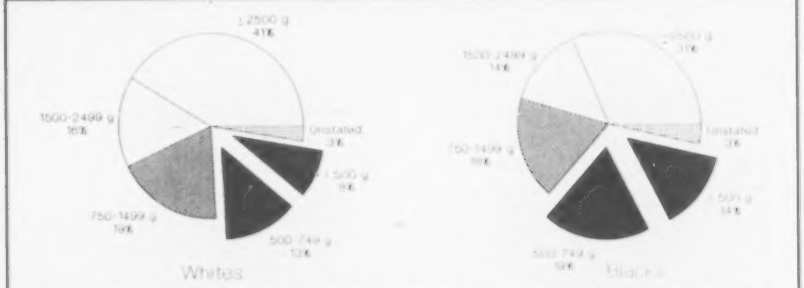


FIGURE 1—Attributable proportions of infant deaths within each race for the US 1983 birth cohort.

Source.—National Center for Health Statistics, National Vital Statistics System (unpublished data).

from the 0.3% of births at these lowest weights (14.7% from White births, 9.2% from Black births). Infants weighing 500 to 749 g had a mortality rate more than three times the rate of those weighing 750 to 1499 g; those weighing under 500 g had nearly four times the rate. Mortality rates with weight not stated indicate a bias toward the smallest infants.

Higher birth-weight-specific mortality rates for Blacks occurred only at normal weights (RR = 1.66) and unknown weights (RR = 2.16). Relative risks of death for Black infants compared with White infants were 0.85 at 1500 to 2499 g, 0.77 at 750 to 1499 g, 0.93 at 500 to 749 g, and 0.98 at under 500 g. Relative risks indicate better survival only for those Black infants weighing between 750 and 1499 g and 1500 and 2499 g.

The attributable proportion of Black deaths from births weighing less than 750 g was 33%, compared with 21% for White deaths (Figure 1). Infants born weighing 750 to 1499 g contributed similar proportions of deaths within each race. Because there were 5.3 times more White births than Black births weighing 2500 g or more, compared with only 1.5 times more White births weighing less than 500 g (Table 1), a much larger fraction of Black deaths resulted from extremely low weights.

The risk of death was highest in the early neonatal period (from 0 to 6 days of life), especially for VLBW infants (Table 2). The early neonatal mortality rate increased exponentially as birth weight decreased by 250-g intervals. The late neonatal and postneonatal mortality rates also increased as birth weight decreased. The small numbers of late neonatal and postneonatal deaths reported for infants weighing less than 500 g may have produced unreliable rates. Survival decreased rapidly as birth weight decreased.

Births weighing less than 750 g resulted in 41% of all early neonatal deaths (Table 3). Of those, 18% were attributable to births weighing less than 500 g. Deaths of the smallest infants contributed to more early neonatal mortality among Blacks than among Whites (Figure 2). Twenty-five percent of Black and 15% of White early neonatal deaths were due to infants weighing <500 grams at birth. Infants weighing less than 750 g contributed 53.2% of Black deaths and 36.4% of White deaths occurring before the first week of life.

Discussion

Changes have occurred both in the reporting of vital statistics on extremely premature infants and in clinical practice affecting fetal and infant survival. Although there are problems with the reporting of deaths among the very smallest infants,[17–19] there is also an increasing number of live-born deliveries that previously might have been fetal deaths.[20–22] Although low-weight births are better reported in the United States than in most developed countries,[23] new initiatives are still needed to improve data quality.[24–26]

Different interpretations of the World Health Organization's definitions of live birth and fetal death in borderline situations affect the reporting of these occurrences in the 500-g range.[27–29] Comparison is complicated by different underlying distributions of extremely low-weight births among races as well as among states and countries.[3,30–34] Adjusting for differences in birth-weight distributions would improve precision in analyses of differences by race or region in infant mortality.[35–37]

Several reports have emphasized that reductions in infant mortality and

TABLE 2—Total US Infant Deaths, Birth-Weight-Specific Mortality Rates, and Survivors, by Race According to Age at Death: 1983 Birth Cohort

	Number of Infant Deaths (1 year)	Mortality Rates					
		Infant (1 year)	Total Neonatal (<28 days)	Early Neonatal (<7 days)	Late Neonatal (7–27 days)	Postneonatal (28–364 days)	Number Surviving[a]
Total[b]	39 683	10.9	7.1	6.0	1.1	3.8	3 599 430
White	27 094	9.3	6.1	5.1	1.0	3.2	2 877 287
Black	11 087	18.9	12.3	10.5	1.8	6.6	574 998
1250–1499 g[b]	1378	114.4	81.9	66.1	15.8	32.5	10 671
White	1026	130.3	96.9	79.2	17.6	33.4	6851
Black	308	82.1	50.1	38.9	11.2	32.0	3442
1000–1249 g[b]	2145	210.4	163.7	130.1	33.7	46.7	8048
White	1545	239.3	195.3	154.1	41.2	44.0	4911
Black	526	155.0	104.0	84.0	20.0	51.0	2867
750–999 g[b]	3762	417.8	341.0	287.1	53.9	76.9	5242
White	2476	442.7	377.4	320.8	56.7	65.3	3117
Black	1167	370.9	275.6	226.6	49.0	95.4	1979
500–749[b]	5815	765.7	712.5	660.0	52.5	53.2	1779
White	3549	790.1	746.4	698.8	47.6	43.6	943
Black	2114	735.6	670.1	609.6	60.5	65.4	760
<500 g[b]	3937	885.9	881.2	875.1	6.1	4.7	507
White	2295	890.6	886.3	877.8	8.5	4.3	282
Black	1534	877.1	871.4	869.1	2.3	5.7	215
Weight not stated[c]	1163	216.1	203.6	193.0	10.6	12.4	4220
White	732	176.5	164.9	154.8	10.1	11.6	3415
Black	388	381.9	366.1	353.3	12.8	15.7	628

Source.—National Center for Health Statistics, National Vital Statistics System (unpublished data).
[a]Infants surviving the first year of life; includes infants who may have died but for whom no death certificate was issued or could be linked.
[b]Includes races other than White and Black and "not stated" weights.
[c]Deaths for which birth weight was not stated on the linked birth certificate.

Black/White differentials will require interventions affecting the lowest weight categories.[3,9,22,38,39] Birth-weight-specific data improve the ability to evaluate such interventions.[40–41] If interventions had prevented deaths of infants weighing less than 500 g in 1983, the IMR of 10.9 per 1000 live births would have been 9.8 (9.3 reduced to 8.5 for Whites, and 18.9 reduced to 16.3 for Blacks). Prevention of deaths among infants weighing less than 750 g would have lowered the IMR to 8.3 (7.3 for Whites and 12.8 for Blacks). Finally, the Black/White mortality differential would have been reduced by 25%. □

TABLE 3—Number and Percent of Total US Early Neonatal Deaths[a] from Births Weighing less than 500 g and less than 750 g: 1983 Birth Cohort

	Neonatal Deaths	Percent			
		Total United States	White	Black	Other
All weights	21 689	100.0	68.5	28.4	3.1
Total <750 g	8901	41.0	24.9	15.1	1.0
500–749 g	5012	23.1	14.5	8.1	0.5
<500 g	3889	17.9	10.4	7.0	0.3

Source.—National Center for Health Statistics, National Vital Statistics System (unpublished data).
[a]Deaths in less than 7 days.

References

1. Hoffman HJ, Bergsjo P, Denman DW. Trends in birth weight-specific perinatal mortality rates: 1970–83. In: *Proceedings of the International Collaborative Effort on Perinatal and Infant Mortality.* Vol 2. Hyattsville, Md: US Dept of Health and Human Services; 1988;3:51–55.
2. Pharoah POD, Alberman ED. Mortality of low birthweight infants in England and Wales, 1953 to 1979. *Arch Dis Child.* 1981;56:86–89.
3. Williams RL, Chen PN. Identifying the sources of the recent decline in perinatal mortality rates in California. *N Engl J Med.* 1982;306:207–214.
4. French FE, Bierman JM. Probabilities of fetal mortality. *Public Health Rep.* 1962;77:835–847.
5. Kline J, Stein Z, Susser M. *Conception to Birth: Epidemiology of Prenatal Development.* New York, NY: Oxford University Press; 1989:59.
6. Wilcox AJ. Surveillance of pregnancy loss in human populations. *Am J Ind Med.* 1983;4:285–291.
7. Moore KL. *The Developing Human: Clinically Oriented Embryology.* 4th ed. Philadelphia, Pa: WB Saunders; 1988:93.
8. Kramer MS. Intrauterine growth and gestational duration determinants. *Pediatrics.* 1987;80:502–511.
9. National Center for Health Statistics. *Health, United States: 1989.* Hyattsville, Md: Public Health Service; 1990. US Dept of Health and Human Services publication PHS 90-1232.
10. Chase H. Infant mortality and weight at birth: 1960 United States birth cohort. *Am J Public Health.* 1969;59:1618–1628.
11. Hogue CJR, Buehler JW, Strauss LT, Smith JC. Overview of the National Infant Mortality Surveillance (NIMS) project: design, methods and results. *Public Health Rep.* 1987;102:126–137.
12. National Center for Health Statistics. Public use data tape documentation. Linked birth/infant death data set: 1983 birth cohort. Hyattsville, Md: US Dept of Health and Human Services; January 1989.
13. Prager K. Evaluation of 1983–85 state and national linked files. In: *Proceedings of the Second Evaluation Meeting of the NCHS Linked Birth and Infant Death Record Project, IV:1–22.* November 13–14, 1989; Houston, Tex.
14. Hoffman HJ, Malloy MH, Stockbauer JW.

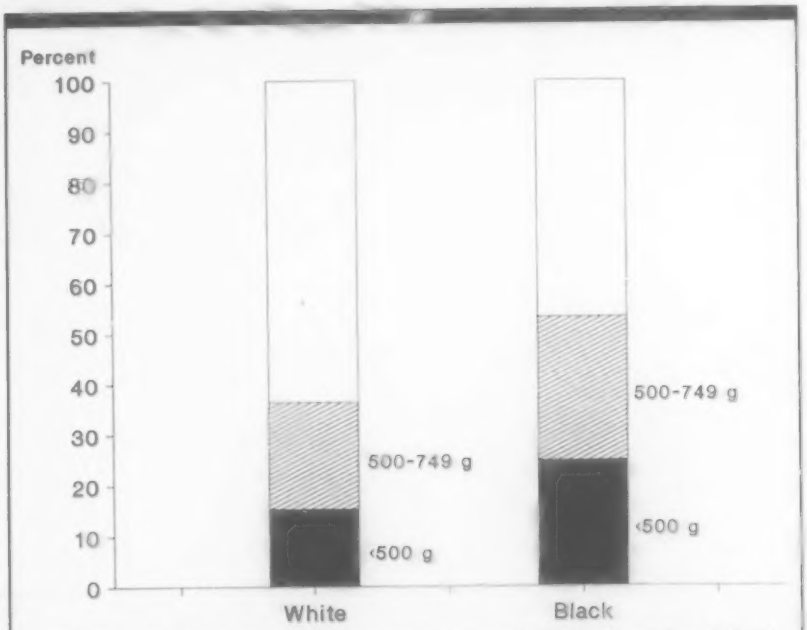


FIGURE 2—Attributable proportions of early neonatal mortality (less than 7 days) within each race from births at 500–749 g and at less than 500 g for the US 1983 birth cohort.

Source.—National Center for Health Statistics, National Vital Statistics System (unpublished data).

Misclassification of preterm births in vital statistics data. *Paediatr Perinat Epidemiol.* 1991;5:A16–A17. Abstract.

15. National Center for Health Statistics. A method of imputing length of gestation on birth certificates. *Vital and Health Statistics.* Series 2, No. 93. Washington, DC: US Govt Printing Office; 1982. US Dept of Health and Human Services publication PHS 82–1367.
16. Fleiss JL. *Statistical Methods for Rates and Proportions.* 2nd ed. New York, NY: John Wiley & Sons; 1981.
17. Goldhaber MK. Fetal death ratios in a prospective study compared to state fetal death certificate reporting. *Am J Public Health.* 1989;79:1268–1270.
18. McCarthy BJ, Terry J, Rochat RW, et al. The underregistration of neonatal deaths, Georgia 1974–1977. *Am J Public Health.* 1980;70:977–982.
19. Task Force on Infant Mortality. Statement on infant mortality. *Pediatrics.* 1986;78: 1155–1160.
20. Hack M, Fanaroff AA. Changes in the delivery room care of the extremely small infant (<750 g): effects of morbidity and outcome. *N Engl J Med.* 1986;314:660–664.
21. US Congress, Office of Technology Assessment. *Neonatal Intensive Care for Low Birthweight Infants: Costs and Effectiveness.* Washington, DC: US Govt Printing Office; December 1987. Publication OTA–HCS–38.
22. US Congress, Office of Technology Assessment. *Healthy Children: Investing in the Future.* Washington, DC: US Govt Printing Office; 1988:45–46. Publication OTA–H–345.
23. European Community Working Group on the Very Low Birthweight Infant. European Community collaborative study of outcome of pregnancy between 22 and 28 weeks' gestation. *Lancet.* 1990;336:782–784.
24. Patterson JE. Assessing the quality of vital statistics. *Am J Public Health.* 1980;70: 944–945.
25. Kleinman JC. Underreporting of infant deaths: then and now. *Am J Public Health.* 1986;76:365–366.
26. Davidson EC. A strategy to reduce infant mortality. *Obstet Gynecol.* 1991;79:1–5.
27. National Center for Health Statistics. *State Definitions and Reporting Requirements for Live Births, Fetal Deaths, and Induced Terminations of Pregnancy.* Hyattsville, Md: US Govt Printing Office; 1981. US Dept of Health and Human Services publication 81–1119.
28. American Academy of Pediatrics and American College of Obstetricians and Gynecologists. Standard terminology for reporting of reproductive health statistics in the United States. *Public Health Rep.* 1988;103:464–471.
29. World Health Organization. *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, based on the recommendations of the Ninth Revision Conference, 1975.* Geneva, Switzerland: World Health Organization; 1977:763–769.
30. Williams RL, Creasy RK, Cunningham GC, et al. Fetal growth and perinatal viability in California. *Obstet Gynecol.* 1982;59:624–632.
31. Goldenberg RL, Humphrey JL, Hale CB, et al. Neonatal deaths in Alabama, 1970–1980: an analysis of birth-weight and race-specific neonatal mortality rates. *Am J Obstet Gynecol.* 1983;145:545–552.
32. Alberman E, Bergsjo P, Cole S, et al. International collaborative effort (ICE) on birthweight; plurality, and perinatal and infant mortality: I. methods of data collection and analysis. *Acta Obstet Gynecol Scand.* 1989;68:5–10.
33. Pratt MW. The changing maternal and child health population: demographic parameters. In: Wallace HM, Ryan GM, Oglesby AC, eds. *Maternal and Child Health Practices.* 3rd ed. Oakland, Calif: Third Party Publishing Co; 1988:68.
34. Lee K, Paneth N, Gartner LM, Pearlman M. The very low-birth-rate: principal predictor of neonatal mortality in industrialized populations. *J Pediatr.* 1980;97:759–764.
35. Hoffman HJ, Meirik O, Bakketeig LS. Methodological considerations in the analysis of perinatal mortality rates. In: Bracken MB, ed. *Perinatal Epidemiology.* New York, NY: Oxford University Press; 1984:504.
36. Wilcox A, Russell I. Why small Black infants have a lower mortality rate than small White infants: the case for population-specific standards for birth weight. *J Pediatr.* 1990;116:7–10.
37. Kleinman JC. Methodological issues in the analysis of vital statistics. In: Kiely M, ed. *Reproductive and Perinatal Epidemiology.* Boca Raton, Fla: CRC Press; 1991:447.
38. Yankauer A. What infant mortality tells us. *Am J Public Health.* 1990;80:653–654.
39. Joyce T. The dramatic increase in the rate of low birthweight in New York City: an aggregate time-series analysis. *Am J Public Health.* 1990;80:682–684.
40. Kleigman RM, Rottman CJ, Behrman RE. Strategies for the prevention of low birth weight. *Am J Obstet Gynecol.* 1990; 162:1073–1083.
41. Public Health Service Expert Panel on the Content of Prenatal Care. *Caring for Our Future: The Content of Prenatal Care.* Washington, DC: US Dept of Health and Human Services; 1989.

An Epidemiological Comparison of Mexican-American and White Non-Hispanic 8th- and 12th-Grade Students' Substance Use

Ernest L. Chavez, PhD, and Randall C. Swaim, PhD

ABSTRACT

Surveys of drug and alcohol use were conducted with a national probability sample of 8th- and 12th-grade Mexican-American and White non-Hispanic youth. The drug and alcohol epidemiology is representative of US Mexican-American youth residing in communities with populations of 2500 or more, at least 10% of whom are Mexican Americans. Mexican-American 8th graders reported generally higher rates of use than White non-Hispanics based on lifetime prevalence and use in last month. They also reported higher frequency of high-risk drug behaviors. The pattern was reversed among 12th-grade students. The impact of differential school dropout rates is discussed as a probable cause for this reversal. (*Am J Public Health.* 1992;82:445–447)

Introduction

Epidemiological studies of Hispanic substance use have focused on use in specific states[1] or specific communities,[2,3] or on use among youth in Puerto Rico.[4,5] In recent reviews of Hispanic drug use, Chavez and Swaim[6] and De La Rosa and colleagues[7] conclude that there are important differences in the drug use of subgroups of Hispanics and that there are substantial gaps in our knowledge of both the epidemiology and etiology of Hispanic substance use. The current study addresses one of the important gaps by presenting normative data on substance use among 8th- and 12th-grade Mexican-American youth. Mexican-American (MA) youth are compared with White non-Hispanic (WnH) youth residing in the same communities.

Method

Sampling Procedure

A representative national sample of 8th- and 12th-grade Mexican-American students was selected from an initial sampling frame of all US communities with populations of 2500 or higher, at least 10% of which was Mexican American (based on the 1980 US Census). Communities were chosen at random throughout the range of population densities so that the sample represents urban, midsized, and rural communities.

Once a community was chosen, individual schools within the community were selected with probabilities equal to the proportion of Mexican Americans within each school. A total of 26 middle schools and 25 high schools were surveyed over the 3-year project.

Subjects

All available students within the class grade in each school were surveyed. Those not surveyed were absentees or refused to participate (less than 1%). Eighty-two percent of all enrolled students completed surveys. This amounted to 2530 8th-grade and 1837 12th-grade MA students and 1547 8th-grade and 2243 12th-grade WnH students attending the same schools. Data on enrollment by ethnicity were not available from all schools, so calculation of the percentage surveyed by ethnic group was not possible. For 8th-grade students, 48.7% were female; for 12th-grade students, 51.0% were female.

The Survey

Students completed the American Drug and Alcohol Survey™, a self-administered survey that measures current and lifetime use of 21 substances. This instrument has been administered to more than 300 000 students, and much of its development has been with minority adolescents. It has been demonstrated to be highly reliable.[8,9]

Results

Lifetime Prevalence

Lifetime prevalence provides a comparative measure between populations of the degree of exposure to an individual drug and of the percentage of individuals who have been willing to at least try the substance. Lifetime prevalence rates for 8th- and 12th-grade students are in Table 1. MA 8th graders reported significantly higher lifetime prevalence rates for marijuana, cocaine or crack, crack only, heroin, steroids, and cigarettes. WnH 8th graders reported higher rates for alcohol, stay-awake pills, diet pills, and smokeless tobacco. WnH 12th graders reported significantly higher lifetime prevalence across a wide range of substances includ-

The authors are with the Tri-Ethnic Center for Prevention Research, Department of Psychology, Colorado State University in Fort Collins.

Requests for reprints should be sent to Ernest L. Chavez, PhD, Department of Psychology, Colorado State University, Fort Collins, CO 80523.

This paper was submitted to the Journal December 4, 1990, and accepted with revisions September 16, 1991.

TABLE 1—Lifetime Prevalence (Percentage of Those Who Ever Tried)

	8th Grade		12th Grade	
	WnH (n = 1547)	MA (n = 2530)	WnH (n = 2243)	MA (n = 1837)
Alcohol	68.6[a]	60.8[a]	93.5[a]	84.5[a]
Drunk	28.5	27.5	74.8[a]	57.6[a]
Marijuana	15.8[a]	24.2[a]	49.7[a]	44.8[a]
Inhalants	15.9	16.5	15.1[a]	12.1[a]
Amyl/butyl nitrite	3.9	4.8	10.3[a]	5.4[a]
Uppers	7.0	8.0	20.2[a]	13.8[a]
Legal stimulants	7.9	9.0	22.7[a]	11.5[a]
Cocaine or crack	3.7[a]	6.3[a]	15.0	13.2
Crack	2.7[a]	5.4[a]	5.6	6.1
Tranquilizers	1.8	2.1	3.9[a]	2.7[a]
Downers	4.0	5.0	5.8	5.0
LSD	5.1	5.6	12.8[a]	7.3[a]
Other psychedelics	4.1	3.9	9.6[a]	4.0[a]
PCP	3.7	4.9	2.8	2.9
Heroin	3.3[a]	4.9[a]	1.8	2.1
Other narcotics	5.3	6.7	11.6[a]	7.3[a]
Crystal meth.	2.3	2.0	8.1	6.6
Stay-awake pills	17.9[a]	11.2[a]	47.3[a]	18.7[a]
Diet pills	11.0[a]	7.1[a]	12.6[a]	19.6[a]
Steroids	0.8[a]	2.9[a]	2.2[a]	1.3[a]
Cigarettes	45.7[a]	49.6[a]	62.5[a]	57.7[a]
Smokeless tobacco	24.4[a]	13.8[a]	33.6[a]	20.3[a]

Note.—WnH = White non-Hispanic; MA = Mexican American.
[a]Indicates significant differences at $P < .05$.

TABLE 2—Current Use (Percentage of Those Who Used in Last Month)

	8th Grade		12th Grade	
Substance	WnH (n = 1547)	MA (n = 2530)	WnH (n = 2243)	MA (n = 1837)
Alcohol	28.5	29.1	60.3[a]	51.2[a]
Drunk	10.3	11.8	35.5[a]	23.0[a]
Marijuana	6.1[a]	11.2[a]	17.9[a]	13.6[a]
Inhalants	5.9[a]	7.5[a]	2.2	2.1
Amyl/butyl nitrate	0.8[a]	1.5[a]	1.3	0.9
Uppers	2.6	3.5	4.9[a]	3.0[a]
Cocaine or crack	2.1	3.1	6.3	6.1
Crack	1.1[a]	2.4[a]	1.9	1.9
Tranquilizers	1.4	1.7	1.3	1.2
Downers	1.7	2.6	1.6	1.3
LSD	2.3	2.9	3.6[a]	2.5[a]
Other psychedelics	1.1	1.6	1.5	0.9
PCP	1.6	2.5	0.6	0.9
Heroin	1.8	2.5	0.7	0.9
Other narcotics	1.3	2.1	1.3	1.7
Cigarettes (half pack+ a day)	2.7	2.1	9.4[a]	2.0[a]
Smokeless Tobacco (6+ times a day)	2.3[a]	1.1[a]	4.4[a]	0.9[a]

Note.—WnH = White non-Hispanic; MA = Mexican American.
[a]Indicates significant differences at $P < .05$.

ing alcohol, ever drunk, marijuana, inhalants, nitrites, uppers, legal stimulants, tranquilizers, LSD, other psychedelics, other narcotics, stay-awake pills, steroids, cigarettes, and smokeless tobacco.

Use in Last Month

The results for last 30-day prevalence are in Table 2. MA 8th graders reported significantly higher rates of use than WnH 8th graders for marijuana, inhalants, nitrites, and crack. Among 12th-grade students, rates were higher among WnH youth for nearly all substances and were significantly so for alcohol, ever drunk, marijuana, uppers, LSD, cigarettes, and smokeless tobacco.

Some youth who use drugs and alcohol will engage in high-risk drug behaviors that are usually part of a pattern associated with heavy substance involvement and may include exposure to either increased health risks due to the use of needles or accident risks due to heavy intoxication. The percentage of youth reporting high-risk substance behaviors is presented in Table 3.

MA 8th graders reported significantly higher rates of high-risk behaviors across a number of categories, including using substances when alone, sharing a needle, combining use of alcohol and marijuana, and using two other drugs at once. Among 12th graders, WnHs reported higher levels of high-risk behaviors in nearly all categories, and significantly in using substances when alone, staying high for several hours, and combining use of multiple substances.

Discussion

These results provide normative rates of drug use epidemiology for Mexican-American youth residing in communities with populations of 2500 or more. These data also supply baseline rates so that comparisons can be made with other racial groups, other Hispanic subgroups, and subgroups of Mexican-American youth (e.g., barrio youth).

There was a general trend among 8th-grade students toward higher use by MAs. The 12th-grade results reveal a general trend toward higher use among WnHs. This shift was also present for rates of high-risk drug behaviors. Several hypotheses may explain these reversals of drug-use rates between the 8th and 12th graders.

First, these differences may be due to cohort effects. But given the gradual decline in adolescent substance use over the past several years, this would mean that whatever factors are leading to this reduction are working selectively on WnH youth and not on MA youth. This scenario is not likely.

A more likely explanation is the dropout effect. Rates of drug use are generally higher for dropouts, and the dropout rate is significantly higher for Hispanics. Those MA youth who are most likely to use and abuse drugs are, therefore, absent from the 12th-grade sample. Chavez et al.[10] reported preliminary results showing that

MA dropouts have higher lifetime prevalence rates than WnH dropouts for many drugs, and that these rates are higher than those reported by either MA youth doing well in school or MA youth who are "educationally at risk." Any attempt to estimate the overall drug use rate of adolescents that does not take into account the dropout problem may be seriously underestimating the extent of the drug use problem within this age group. Further epidemiological research on drug use with various racial groups that includes drug use data on dropouts is needed to determine how dropping out of school affects substance use estimates. □

TABLE 3—Lifetime High-Risk Substance Behaviors

	8th Grade		12th Grade	
	WnH	MA	WnH	MA
High-Risk Behavior	(n = 1547)	(n = 2530)	(n = 2243)	(n = 1837)
Got drunk alone	9.3[a]	11.6[a]	23.0[a]	14.8[a]
Used marijuana alone	5.6[a]	11.7[a]	17.1[a]	13.2[a]
Used other drug alone	6.1[a]	8.3[a]	11.8[a]	9.4[a]
Needle/cocaine	2.0	2.6	1.3	1.1
Needle/heroin	2.3	2.8	1.3	1.2
Shared a needle	1.6[a]	2.7[a]	0.9	1.1
Stayed high 7+ hours	4.4	5.6	16.1[a]	9.9[a]
Used alcohol and marijuana	8.5[a]	12.9[a]	36.7[a]	26.9[a]
Used alcohol and other drug	7.0	8.2	20.7[a]	14.0[a]
Used two drugs at same time	5.4[a]	7.5[a]	16.5[a]	12.1[a]

Note.—WnH = White non-Hispanic; MA = Mexican American.
[a]Indicates significant differences at $P < .05$.

Acknowledgments

Preparation of this article was funded, in part, by the Tri-Ethnic Center for Prevention Research, NIDA grant #P50DA07074, and NIDA grant #R01DA04000.

The authors wish to acknowledge E. R. Oetting and F. Beauvais for their helpful reviews of earlier versions of this article.

References

1. Fredlund EV, Spence RT, Maxwell JC. *Substance Use among Students in Texas Secondary Schools: 1988.* Austin, Tex: Texas Commission on Alcohol and Drug Use; 1989.
2. Padilla ER, Padilla AM, Morales A, Olmedo EL, Ramirez R. Inhalant, marijuana, and alcohol abuse among barrio children and adolescents. *Int J Addict.* 1979;14:943–964.
3. Perez R, Padilla AM, Ramirez A, Ramirez R, Rodriguez M. Correlates and changes over time in drug and alcohol use within a barrio population. *Am J Community Psychol.* 1980;6:621–636.
4. Robles R, Martinez R, Muscoso M. Drug use among public and private secondary school students in Puerto Rico. *Int J Addict.* 1979;14:243–258.
5. Velez CN, Ungemack JA. Drug use among Puerto Rican youth: an exploration of generational status differences. *Soc Sci Med.* 1989;29:779–789.
6. Chavez EL, Swaim RC. Hispanic substance use: problems in epidemiology. *Drugs and Soc.* In press.
7. De La Rosa M, Khalsa JH, Rouse BA. Hispanics and illicit drug use: a review of recent findings. *Int J Addict.* 1990; 25(6):665–691.
8. Oetting ER, Beauvais F. Adolescent drug use: findings of national and local surveys. *J Consult Clin Psychol.* 1990;58:385–394.
9. Oetting ER, Beauvais F, Edwards RW, Waters M. *The Clinical Drug Abuse Scale.* Book 2, *Instrument Development, Reliability, and Validity.* Fort Collins, Colo: Rocky Mountain Behavioral Sciences Institute, Inc; 1984.
10. Chavez EL, Edwards RW, Oetting ER. Mexican-American and White-American dropouts: drug use, health, and violence. *Public Health Rep.* 1989;104:594–604.

Analyzing the Effect of Prenatal Care on Pregnancy Outcome: A Conditional Approach

Michael H. Malloy, MD, MS, Tzu-Cheg Kao, PhD, and Y. J. Lee, PhD

ABSTRACT

We used vital statistics data to examine the effect of the time of entry into prenatal care relative to the gestational age of delivery on pregnancy outcome. Early entry into prenatal care was associated with better outcomes only in women who delivered at 37 through 42 weeks of gestation. We are concerned that the evaluation of the effect of prenatal care using vital statistics data may be questionable because of the problems inherent in the data. (*Am J Public Health.* 1992;82:448–450)

Introduction

Vital statistics data is a readily available source of information on prenatal care usage in the United States. How to examine vital statistics and other types of data for the effect of prenatal care on pregnancy outcome has generated a number of different methodologies.[1–4] One of the difficulties encountered in using such data has been the problem of overcoming the preterm bias effect. The preterm bias effect is associated with observing better outcomes in women who receive prenatal care than in women who receive no care. This artifact occurs because women receiving no care may deliver prematurely and may not have an opportunity to enroll in prenatal care.[5] By conditioning the effect of the age of entry into prenatal care on the gestational age of delivery, however, the preterm bias effect can be avoided. Nonetheless, other problems encountered in vital statistics data may invalidate the results of the conditional approach.

In this study, we demonstrate the effect that the possible misclassification of gestational age in vital statistics data has on the interpretation of the conditional approach and discuss the problematic nature of the evaluation of prenatal care.

Methods

The data for this analysis were obtained from linked birth and death certificates from Missouri for 1980 through 1984. The variables used in the analysis were the gestational age at which prenatal care was initiated (self-report and limited to the first 36 weeks of pregnancy), the gestational age at delivery (based on last menstrual period), the outcome (stillbirth, infant death, or alive at 1 year), race (Black or non-Black), birth number (multiple vs single), parity, maternal age, and maternal education. We used only those records that involved gestational ages of greater than 19 weeks and less than 51 weeks. Of the 387 173 records available, 375 858 (97.1%) were used for the unadjusted analyses and 374 244 (96.7%) were used for the adjusted analyses.

The association of the gestational age of entry into prenatal care with pregnancy outcome was examined relative to the gestational age of delivery. Outcomes (stillbirths or infant deaths) are presented as death rates (percentage dead). We calculated odds ratios for stillbirths or infant deaths by logistic regression models adjusted for race (Black vs non-Black), multiple births (multiple vs single), (0 parity and parity greater than 3 compared with parity 1 through 3), maternal age (less than 18 and greater than 34 compared with 18 through 34 years), and maternal education (less than 12 compared with greater than 12 years). Mantel-Haenszel tests for linear trends were carried out for mortality rates over the gestational ages of entry into prenatal care.[6] Furthermore, we adjusted our results by declaring newborns that exceeded the 90th percentile birth-weight standards for gestational age reported by Arbuckle[7] as misclassified. All computations were carried out using SAS®.[8]

Results

By conditional analysis, the risk (percentage of mortality) of stillbirth or infant death for those women who delivered at 25 weeks to 28 weeks of gestation and who entered prenatal care between 1 and 4 weeks was greater than the risk of still-

At the time of this study both Michael H. Malloy and Y. J. Lee were with the Division of Prevention Research, National Institute of Child Health and Human Development, Bethesda, Maryland. Michael H. Malloy is now with the Department of Pediatrics of the University of Texas Medical Branch. Tzu-Cheg Kao is with the Department of Preventive Medicine and Biometrics, Uniform Services University of the Health Sciences, Bethesda, Maryland.

Requests for reprints should be sent to Michael H. Malloy, MD, MS, Department of Pediatrics, 6.104 Waverly-Suite Pavillion, University of Texas, Medical Branch, Galveston, TX 77550.

This paper was submitted to the journal February 5, 1991, and accepted with revisions July 10, 1991.

TABLE 1—Percentage of Mortality (Stillbirth or Infant Death) Conditional on Gestational Age of Entry into Prenatal Care and Gestational Age of Delivery

Gestational Age at Entry into Prenatal Care, wks	Gestational Age of Delivery, wks					
	20–24	25–28	29–32	33–36	37–42	43–50
1–4	85.10	50.93	15.69	3.58	0.60	0.82
5–8	86.45	55.36	19.68	4.46	0.72	0.78
9–12	74.46	48.50	17.34	4.12	0.89	0.97
13–16	62.60	41.60	14.73	3.43	1.03	1.16
17–20	45.12	33.93	10.59	4.11	1.27	1.77
21–24	...	31.51	9.61	4.12	1.63	1.68
25–28	...	...	8.63	4.23	1.10	1.48
29–32	...	...	...	4.00	1.61	2.15
33–36	...	...	...	...	1.74	1.25
No care	68.10	35.16	21.08	8.44	2.96	1.43
χ^2	74.3	33.3	25.3	0.018	168.9	35.3
P value[a]	<.001	<.001	<.001	.89	<.001	<.001
Total deliveries	1233	2129	5132	23 892	299 690	43 793

[a] Mantel-Haenszel test for linear trends in mortality rates excluding those with no care.[6]

TABLE 2—Percentage of Mortality (Stillbirth or Infant Death) Conditional on Gestational Age of Entry into Prenatal Care and Gestational Age of Delivery[a]

Gestational Age at Entry into Prenatal Care, wks	Gestational Age of Delivery, wks					
	20–24	25–28	29–32	33–36	37–42	43–50[b]
1–4	91.67	56.67	19.58	3.91	0.58	0.82
5–8	94.66	61.92	23.92	4.76	0.71	0.78
9–12	91.43	62.21	23.00	4.37	0.85	0.97
13–16	91.89	54.91	22.16	3.65	0.96	1.16
17–20	93.55	54.44	18.07	4.51	1.19	1.77
21–24	...	51.35	16.10	4.15	1.55	1.68
25–28	...	...	14.06	4.43	1.02	1.48
29–32	...	...	...	4.76	1.42	2.15
33–36	...	...	...	...	1.68	1.25
No care	84.29	47.50	26.00	8.57	2.97	1.43
χ^2	0.50	1.23	3.07	0.051	226.0	3.53
P value[c]	.822	.268	.080	.821	<.001	<.001
Total deliveries	923	1582	3463	20 509	290 397	43 793

[a] Newborns weighing over the 90th percentile for gestational age, based on the growth standards reported by Arbuckle,[7] were deleted.
[b] Ninetieth percentile weights were not available of gestations greater than 42 weeks, therefore the rates in this table are the same as those presented in Table 1 for outcomes occurring after 42 weeks.
[c] Mantel-Haenszel test for linear trends in mortality rates excluding those with no care.[6]

birth or infant death for women who delivered at 25 to 28 weeks and received no care (Table 1). The adjusted odds ratio was 1.67 (95% confidence interval = 1.03, 2.71). This higher risk of mortality in early entrants delivering before 29 weeks, however, appears to be an artifact of the misclassification of the gestational age of the survivors. This is apparent when newborns with weights greater than the 90th percentile for gestational age were not included in the analysis (Table 2). The declining mortality rates as the gestational age of entry into prenatal care increased observed in Table 1 are not apparent in Table 2.

For women delivering after 36 weeks of gestation, early entry into prenatal care was a marker for optimal pregnancy outcome. The risk for stillbirth or infant death increased progressively for those delivering at 37 through 42 weeks from a low of 0.60% for those who entered prenatal care at 1 through 4 weeks to a high of 2.96% for those who received no care (Table 1). The results for this group in Table 1 are similar to the results reported in Table 2.

Discussion

As Tyson et al. state, "the value of prenatal care has been difficult to establish."[4] When the time of entry into prenatal care is assessed relative to the time of delivery, several interpretations about the association of prenatal care with pregnancy outcome become available. For example, it appears that in women who delivered at less than 29 weeks and who entered prenatal care in the first trimester, there were more adverse pregnancy outcomes than in those women who delivered at this time and were not in prenatal care. We believe this is an artifact. Kramer et al. have reported a systematic underestimation of gestational age in infants born before 37 weeks when the gestational age is based solely on the last menstrual period.[9] When infants weighing over the 90th percentile for their gestational age were eliminated from the analysis, there were no significant trends in mortality rates over the gestational age of entry into prenatal care except for those women delivering

between 37 and 42 weeks. Whether prenatal care for these women is responsible for the better outcome of the early entrants cannot be determined. Early entry may reflect as much on the healthy behavioral characteristics of these women as it does on any effect of prenatal care.

The evaluation of prenatal care is a complex problem. The incorrect assignment of gestational age affects significantly the results of any perinatal analysis and continues to be a major problem for which no solution is readily available. As a result, there is no ideal model available to assess the association between prenatal care and pregnancy outcome. Because of the substantial problems associated with observational data, the use of randomized trials of different modes of care deserves increased consideration as the optimal method for gaining a greater understanding of the efficacy of prenatal care. □

Acknowledgments

We gratefully acknowledge the provision of the Missouri vital statistics data by Wayne Schramm and Garland Land.

References

1. Gortmaker SL. The effects of prenatal care upon the health of the newborn. *Am J Public Health.* 1979;69:653–660.
2. Quick JD, Greenlick MR, Roghmann KJ. Prenatal care and pregnancy outcome in an HMO and general population: a multivariate cohort analysis. *Am J Public Health.* 1981;71:381–390.
3. Showstack JA, Budetti PP, Minkler D. Factors associated with birth weight: an exploration of the roles of prenatal care and length of gestation. *Am J Public Health.* 1984;74:1003–1008.
4. Tyson J, Guzick D, Rosenfeld CR, et al. Prenatal care evaluation and cohort analyses. *Pediatrics.* 1990;85:195–204.
5. Terris M, Glasser M. A life table analysis of the relation of prenatal care to prematurity. *Am J Public Health.* 1974;64:869–875.
6. Mantel N. Chi-square tests with one degree of freedom: extensions of the Mantel Haenszel procedure. *J Am Stat Assoc.* 1963;58:690–700.
7. Arbuckle TE, Sherman GJ. An analysis of birth weight by gestational age in Canada. *J Can Med Assoc.* 1989;140:157–165.
8. SAS Institute Inc. *SAS User's Guide: Statistics, Version 5 Edition.* Cary, NC: SAS Institute Inc; 1985.
9. Kramer MS, McClean FH, Boyd ME, Usher RH. The validity of gestational age estimation by menstrual dating in term, preterm, and postterm gestations. *JAMA.* 1988;260:3306–3308.

The Safety of Home Birth: The Farm Study

A. Mark Durand, MD, MPH

ABSTRACT

Pregnancy outcomes of 1707 women, who enrolled for care between 1971 and 1989 with a home birth service run by lay midwives in rural Tennessee, were compared with outcomes from 14 033 physician-attended hospital deliveries derived from the 1980 US National Natality/National Fetal Mortality Survey. Based on rates of perinatal death, of low 5-minute Apgar scores, of a composite index of labor complications, and of use of assisted delivery, the results suggest that, under certain circumstances, home births attended by lay midwives can be accomplished as safely as, and with less intervention than, physician-attended hospital deliveries. (*Am J Public Health.* 1992;82:450–452)

Introduction

Despite a sizable body of literature on the subject, the authors of two recent reviews have concluded that a final judgment regarding the relative safety of home birth still cannot be made.[1,2]

This study compares the outcomes of a cohort of 1707 planned lay midwife–attended home births from the Farm midwifery service in rural Tennessee with the outcomes of a sample of physician-attended hospital births derived from the 1980 US National Natality/National Fetal Mortality Survey (NNS/NFMS). This is the largest comparative study of a cohort of home births yet published.

Previous studies of home birth safety underscore the importance of having explicit comparison groups (rather than of being strictly descriptive),[1] of considering planned and unplanned home births separately,[2–4] of tracking intrapartum hospital transfers of intended home births (and attributing the outcome to the home birth group),[5] and of controlling for the effects of birth weight, type of delivery attendant, and demographic characteristics of the mother.[6,7] All these issues have been addressed in this study.

Standards of prenatal care at the Farm are modeled to the recommendations of the American College of Obstetricians and Gynecologists. Deliveries are conducted without analgesia, however, and great emphasis is placed on meeting the emotional needs of the family. Several family members and friends are commonly in attendance and are encouraged to take an active role in the birth.[8] In addition, the laboring woman is encouraged to stay off her back and remain physically mobile through labor and delivery. In the absence of signs of fetal distress, women are permitted to labor beyond 24 hours, occasionally for 2 to 3 days. They are also encouraged to eat and drink during labor in the belief that this allays maternal exhaustion and the need for operative delivery.

A. Mark Durand is with the Department of Health, Commonwealth of the Northern Marianas Islands.

Requests for reprints should be sent to A. Mark Durand, MD, MPH, Department of Health, Commonwealth of the Northern Marianas Islands, Rota Health Center, Rota, MP 96951.

This paper was submitted to the journal November 11, 1990, and accepted with revisions July 9, 1991.

A hospital with surgical facilities is located about a 20-minute drive from the community.

Methods

A retrospective cohort of 1707 births derived from all pregnancies accepted for care at the Farm between January 1971 and June 1989, except those that aborted spontaneously prior to 28 weeks gestational age and 54 who were lost to follow-up, was studied. The outcomes of all these births were attributed to the Farm group, whether or not delivery was ultimately accomplished at home (Table 1).

Women with preexisting diabetes or hypertension, Rh negative blood with positive antibody screen, weight greater than 135 kg, and hematocrit on intake lower than 28 were considered ineligible for care at the Farm. In addition, before 1985, women with prior cesarean section were also considered ineligible.

The NNS/NFMS is a probability sample of births in the United States in 1980 for which a birth or fetal death certificate was filed. Low birth weights (less than 2500 g) and fetal deaths were deliberately oversampled to allow more detailed study of them.[9] Stillborn infants of less than 28 weeks gestational age (or weight of 1000 g) were not included.

Of the 16 327 births that ultimately constituted the NNS/NFMS, an additional 2294 were deleted from this study to eliminate those that were attended by nonphysicians, those who delivered out-of-hospital, those with no prenatal care, and those in which the mother had prenatal risk factors that were used as exclusion criteria at the Farm (Table 1).

The design effect created by differential sampling and losses due to lack of data across the strata (fetal deaths, live births of low birth weight, other live births) in the NNS/NFMS was compensated for by assigning weighting coefficients (0.0131, 0.4227, and 1.8844, respectively) to the observations in each of the three strata. The coefficients are normalized such that the average coefficient of a single observation is 1.0, which is the coefficient assigned to observations in the Farm group. These coefficients were applied in the analyses described below.

Data for both groups were obtained by retrospective review of medical records.

The principal independent variable of this study is the intended site of delivery. The dependent variables are measures of pregnancy outcome defined as (1) death of the fetus at birth weight greater than 1000 g or at 28 weeks gestational age, or death of a live born infant aged less than 28 days; (2) a 5-minute Apgar score of less than 7; (3) use of forceps, vacuum, or cesarean delivery; and (4) occurrence of excessive bleeding during labor, labor of greater than 24 hours duration, neonatal birth injury, or neonatal respiratory distress. The effect of intended site of delivery on each outcome was examined via multiple logistic regression equations, which also contained terms for potential confounders.

Of these potential confounders, maternal education and number of prenatal visits were treated as continuous variables. Birth order (1, 2 to 4, greater than 4), maternal age (younger than 18, 18 to 34, older than 34), birth weight (less than 2500 g, 2500 to 4000 g, greater than 4000 g), and smoking during pregnancy (yes, no) were categorized.

The power to detect odds ratios of 1.5 or greater for the delivery type and labor complications outcomes exceeds 0.99. For low 5-minute Apgar scores, the corresponding power is 0.75 (0.99 for odds ratio of 2 or greater). For perinatal mortality, this study has a power of 0.62 to detect an odds ratio of 1.5 (0.98 for odds ratio of 2 or greater). All power calculations presume two-sided comparisons.

Results

Demographic characteristics of women in the two groups and transfers to hospital care from the Farm group are summarized in Tables 2 and 3.

No significant differences between the two groups regarding fetal and neonatal death, labor-related complications, or low 5-minute Apgar scores were detected, regardless of whether the comparisons were crude or adjusted for confounders (Table 4). Rates of assisted deliveries were very low in the Farm group (2.11%) compared with the NNS/NFMS sample (26.60%), a difference that is statistically significant at the P = .01 level. The casarean section rate was 1.46% at the Farm vs 16.46% in the NNS/NFMS sample (Table 4).

The causes of death for the 17 deceased neonates and stillbirths in the Farm group are listed in Table 5. There were no maternal deaths in the Farm group.

Discussion

In this study, lay midwife–attended home births appear to have been accomplished with safety comparable to that of conventional births. Furthermore, the proportion of deliveries in which opera-

TABLE 1—Selection of Records for Analysis

Farm group	
Accepted for care	1761
Lost to follow-up	54
Number analyzed	1707
NNS/NFMS	
Number originally sampled	18 006
Insufficient data	1679
Included in NNS/NFMS	16 327
Criteria for present study not met[a]	2294
Number analyzed	14 033

Note. NNS/NFMS = US National Natality/National Fetal Mortality Survey.
[a]Exclusion criteria include no prenatal care, out-of-hospital birth, nonphysician attendant, prepregnancy diabetes or hypertension, anemia (hematocrit lower than 28), weight greater than 135 kg, and Rh negative blood with positive antibody screen.

TABLE 2—Characteristics of 1707 Women Intending Home Birth and of 14 033 Women from the 1980 US NNS/NFMS

Characteristic	Farm Group	NNS/NFMS
Age (%)		
< 18 years	1.3	1.7
18–34 years	94.9	93.6
> 34 years	3.8	4.7
Education (%)		
< 9 years	6.3	3.7
9–12 years	41.7	58.8
> 12 years	52.0	37.5
Unmarried (%)	15.0	18.2
Race (%)[a]		
White	—	82.5
non-White	—	17.5
Birth order (%)		
1	44.7	37.3
2–4	50.7	56.1
> 4	4.6	6.6
Smoking (%)	1.1	12.1
Number of prenatal visits	8.9	10.3
Gestational age of infants (%)		
< 36 weeks	1.5	5.7
36–42 weeks	97.2	89.6
> 42 weeks	1.3	4.7
Birth weight of infants (%)		
< 2500 g	3.8	5.8
2500–4000 g	87.9	83.4
> 4000 g	8.3	10.8

Note. NNS/NFMS = US National Natality/National Fetal Mortality Survey.
[a]An estimated 98% of Farm group mothers are White.

TABLE 3—Transfers to Inpatient Care from the Farm Group

Time of transfer	(n)	(%)
Before start of labor	28	(1.6)
Intrapartum	126	(7.4)
Postpartum[a]	76	(4.5)
Total Referred	230	(13.5)
To hospital during labor without being referred	3	
Delivered at home	1551	(90.1)

[a]Includes infants and mothers.

tive assistance was required was much smaller in the Farm group (though a fraction of this discrepancy may have been due to the fact that, prior to 1985, women who had had previous cesarean deliveries had been excluded from care by Farm midwives).

Stratifying both groups by birth weight or parity, comparing the entire Farm set to the NNS/NFMS subset of births from nonmetropolitan southern US areas or of White mothers only, and breaking the Farm set into groups by year of delivery did not appreciably alter the results (data not shown). Thus, differences in demographic characteristics or changes in technology over time do not appear to explain the findings. Of course, other factors such as diet (many mothers in the Farm group were vegetarians), occupational profile, and psychological differences among those electing home birth could help explain the observed results.

Observer bias could also have distorted the results. (The likelihood of a complication or outcome being noticed and recorded may be different for midwives than for physicians.) However, the results regarding perinatal deaths, in particular, are hard to ascribe to observer bias.

The possibility that some outcomes that were not examined in this study may be worse among those delivering at home and that other groups of attendants may not perform as well as the Farm midwives should also be noted.

Hospital births have a safety advantage in cases in which life-saving technology is immediately required. There is some evidence, however, that elective interventions, which are used more frequently in-hospital, may increase the risk of various adverse outcomes in low-risk women.[10–17]

In addition, it is possible that the unfamiliar setting and the presence of unfamiliar personnel, the limited presence and role of family members, and the restricted freedom of movement of the laboring woman may all create an atmosphere at a hospital birth that undermines self-confidence and encourages passivity on the part of the laboring woman, diminishing her ability to deliver spontaneously. Recent work demonstrating the beneficial effect of supportive female attendants (doulas) during labor, highlights the importance of such intangibles.[18]

The results of this study suggest that, for relatively low-risk pregnancies, home birth with attendance by lay midwives is not necessarily less safe than conventional (hospital-physician) delivery. Support by the medical and legal communities for those electing, and those attending, home birth should not be withheld on the grounds that this option is inherently unsafe. □

TABLE 4—Association between Intended Site of Delivery and Selected Pregnancy Outcomes

Outcome	% Farm group	% NNS/NFMS[a]	Crude RR	Adjusted OR[b]	Adjusted (95% CI)	*P* value
Perinatal death	1.00	1.33	0.75	0.69	(0.38–1.26)	.23
Labor-related complications	6.27	7.29	0.86	0.81	(0.63–1.04)	.09
Bleeding	1.93	1.02	—	—	—	—
labor >24 hours	2.87	2.76	—	—	—	—
Birth injury	0.23	3.34	—	—	—	—
RDS	1.41	3.65	—	—	—	—
Assisted delivery[c]	2.11	26.60	0.08	0.04	(0.03–0.05)	.00
Cesarean section	1.46	16.46	0.09	0.09	(0.08–0.10)	.00
5-minute Apgar<7	1.62	2.40	0.68	0.69	(0.40–1.19)	.18

Note. NNS/NFMS = US National Natality/National Fetal Mortality Survey; RR and OR denote risk and odds ratios; CI = confidence interval; RDS = respiratory distress syndrome.
[a]NNS/NFMS is the reference group.
[b]Adjusted odds ratios were obtained by logistic regression and controlled for maternal age, parity, education, marital status, birth weight, smoking, and number of prenatal visits.
[c]Assisted delivery is use of any of the following: cesarean section, forceps, or vacuum extractor.

TABLE 5—Perinatal Deaths among 1707 Infants of Women Enrolled for Care at the Farm

Category of Death	Number
Lethal congenital anomalies	6
Complications related to prematurity	4
Death in utero before onset of labor	2
Neonatal sepsis	1
Abruption (during labor)	1
Respiratory distress, unknown cause	1
Prolapsed cord	1
Suspected child abuse	1
Total	17

References

1. Committee on Assessing Alternative Birth Settings, Institute of Medicine, Division of Health Sciences Policy and National Research Council, Commission on Life Sciences. Research issues in the assessment of birth settings: report of a study. Washington, DC: National Academy Press; 1982:4.
2. Campbell RA, MacFarlane A. Place of delivery: a review. *Br J Obstet Gynaecol.* 1986;93:675–683.
3. Hinds MW, Bergeson GH, Allen DT. Neonatal outcome in planned vs unplanned out-of-hospital births in Kentucky. *JAMA.* 1985;253:1578–1582.
4. Shy KK, Frost F, Ullom J. Out-of-hospital delivery in Washington State, 1975–1977. *Am J Obstet Gynecol.* 1980;137(5):547–552.
5. Shearer JML. Five-year prospective survey of risk of booking for a home birth in Essex. *Br Med J.* 1985;291(20):1478–1480.
6. Schramm WF, Barnes DE, Bakewell JM. Neonatal mortality in Missouri home births, 1978–1984. *Am J Public Health.* 1987;77(8):930–935.
7. Fedrick J, Butler NR. Intended place of delivery and perinatal outcome. *Br Med J.* 1978;76(1):763–765.
8. Gaskin IM. *Spiritual Midwifery.* Summertown, Tenn: The Publishing Company; 1977.
9. Keppel KG, Heuser RL, Placek PJ, et al. Methods and response characteristics, 1980 NNS and NFMS. *Vital and Health Statistics.* Series 2, no. 100. Hyattsville, Md: National Center for Health Statistics; 1986.
10. Tew M. Do obstetric intranatal interventions make birth safer? *Br J Obstet Gynaecol.* 1986;93:659–674.
11. Klein M, Lloyd F, Redman C, Bull M, Turnbull AC. A comparison of low-risk pregnant women booked for delivery in two systems of care: shared care (consult-

ant) and integrated general practice unit. *Br J Obstet Gynaecol.* 1983;90:118–128.
12. Levine MG, Holroyde J, Woods JR, Siddiqi TA, Scott M, Miodovnik M. Birth trauma: incidence and predisposing factors. *Obstet Gynecol.* 1984;63:792–795.
13. Banta HD, Thacker SB. Assessing the costs and benefits of electronic fetal monitoring. *Obstet Gynecol Surv.* 1979;35:627–642.
14. Haverkamp AD, Orleans M, Langendoerfer S, McFee J, Murphy J, Thompson HE. A controlled trial of the differential effects of intrapartum fetal monitoring. *Am J Obstet Gynecol.* 1979;134:399–412.
15. Brinsden PR, Clark AD. Postpartum hemorrhage after induced and spontaneous labour. *Br Med J.* 1978;2:855–856.
16. Hack M, Fanaroff AA, Klaus MH, Mendelawitz BD, Merkatz IR. Neonatal respiratory distress following elective delivery: a preventable disease? *Am J Obstet Gynecol.* 1976;126:43–47.
17. Baumgarten K. Advantages and disadvantages of low amniotomy. *J Perinat Med.* 1976;89:963–967.
18. Kennell J, Klaus M, McGrath S, Robertson S, Hinkley C. Continuous emotional support during labor in a US hospital: a randomized, controlled trial. *JAMA.* 1991;265(17):2197–2201.

Skilled Nursing Facility Care for Persons with AIDS: Comparison with Other Patients

James H. Swan, PhD, A. E. Benjamin, PhD, and Andrew Brown, MD

ABSTRACT

We compared average daily hours of care for patients with and without acquired immunodeficiency syndrome (AIDS) in a hospital-based skilled nursing facility. Patients with AIDS averaged over an hour more direct nursing care per patient day (6.5 vs 5.4 hours). These findings support other reports of greater nursing time and higher costs of caring for nursing home patients with AIDS. (*Am J Public Health.* 1992;82:453–455)

Introduction

Skilled nursing facility (SNF) care for persons with acquired immunodeficiency syndrome (AIDS) (PWAs) is limited but will likely grow.[1] As the postdiagnosis life expectancy of PWAs increases[2] and greater numbers of PWAs find no other locus of care,[3–5] SNFs should become accepted parts of an AIDS continuum of care.[6–8] AIDS poses major challenges to providers that are inexperienced with PWAs[9,10] and the populations from which they derive,[11] and that are not equipped to deal with special care needs[9] or special precautions and infection controls.[10]

PWAs are generally thought to be costlier than traditional nursing home patients[12,13] because they need different, more complex, and higher volumes of care.[10] Kerschner reports that PWAs in the final months of life require, on the average, 7 hours of nursing home care daily.[14] Adams[15] reports that PWAs require an average of over 5 hours' direct intermediate care facility–level nursing care daily. Swan and Benjamin report an average daily total nursing time of 6.5 hours (5.4 hours direct time) for PWAs in a freestanding AIDS-dedicated SNF[1]; this time approaches the 6.5 to 7.5 hours of care required daily for functionally dependent nursing home patients with complex needs.[16]

Some industry observers have argued that PWAs' need for heavy care has been overstated,[17] but no research has directly examined the relative costs of SNF care for PWAs versus "traditional" nursing home populations. This study begins to fill the gap, using comparative data on AIDS and non-AIDS care in a single hospital-based SNF in California.

Methods

In our study of care for PWAs in two Northern California SNFs with dedicated AIDS beds, a random sample of 29 non-AIDS patients was drawn from a hospital-based SNF between August 18 and December 31, 1987. (The other SNF served PWAs only.) The sample was drawn from two wards that also cared for AIDS patients—an oncology ward and a general ward. These 29 patients averaged 51 days of care per patient, compared with 41 days per patient for the 20 PWAs receiving care in the SNF in the same period.

A major issue was whether nursing hours were greater for PWAs than for others. Nursing hours were measured in two ways: (1) monthly nursing personnel time divided by days of care provided that month, and (2) daily direct plus indirect nursing time per patient attributed to nursing procedures recorded using GRASP,[18]

At the time of this study, all authors were with the Institute for Health Policy Studies, University of California, San Francisco. James H. Swan is now with the Health Care Administration Program, California State University, Long Beach, Calif.

Requests for reprints should be sent to James H. Swan, PhD, Health Care Administration Program, California State University, Long Beach, CA 90840.

This paper was submitted to the journal December 11, 1990, and accepted with revisions June 10, 1991.

TABLE 1—Analysis of Variance: Average Daily Direct Nursing Care Hours for Persons with AIDS (PWAs, n = 20) and Non-AIDS Patients (n = 29) in a Skilled Nursing Facility

	Mean	SEM
PWAs	6.501	0.263
Non-AIDS patients	5.434	0.195
$R^2 = .184$		
Mean square between = 1278.18		
Mean square within = 120.33		
$F(1, 47) = 10.623$, $P < .01$		

Note. Patient days for PWAs = 821; patient days for non-AIDS patients = 1486.

TABLE 2—Analysis of Variance: Average Daily Direct Nursing Care Hours for Persons with AIDS (PWAs) and Non-AIDS Patients in a Skilled Nursing Facility, by Ward

	Mean[a]	SEM
Oncology ward		
PWAs (n = 11)	6.695	0.341
Non-AIDS patients (n = 15)	5.146	0.354
General ward		
PWA (n = 9)	6.212	0.415
Non-patient (n = 14)	5.561	0.341
$R^2 = .215$		
Mean square between = 497.313		
Mean square within = 120.924		
$dF(3, 45) = 4.11$, $P < .05$		

Note. Patient days = 491 for PWAs on the oncology ward, 455 for non-AIDS patients on the oncology ward, 330 for PWAs on the general ward, and 1031 for non-AIDS patients on the general ward.
[a]The difference between mean hours of care for PWAs and mean hours of care for non-AIDS patients on the oncology ward was significant at $P < .01$.

adapted to the SNF. (GRASP is proprietary to FCG Enterprises, Inc, Morgantown, NC.) A comparison showed close agreement between the two methods.[1]

Because patients differ in numbers of days of care and because of a general concern for daily resource use, daily nursing hours are reported. Analysis of variance was employed to test differences in nursing hours.

Findings

PWAs in the hospital-based SNF were more severely ill and functionally debilitated than were PWAs in the community. Two thirds scored 30 or lower on the Karnofsky Performance Status Scale; three fourths had diagnoses of dementia; half were dependent in all activities of daily living; three fourths were discharged by reason of death. Three fourths listed their sexual orientation as homosexual; one third reported intravenous drug use; half were aged 40 or older. Except in age, these PWAs were relatively similar to PWAs nationwide.[19,20]

An analysis of variance was conducted to compare average daily direct nursing time for PWAs and non-AIDS patients (Table 1). PWAs averaged over an hour more direct nursing time (6.5 hours) than did non-AIDS patients (5.4 hours). The latter group still represents high resource use, far greater than the 3.0 total nursing hours defining SNF care in California. This difference is attributable to the SNF's being hospital-based and thus associated with heavy care.

Although the small sample size restricted analytic power, limiting the ability to control for covariates, an analysis of variance was used to compare differences between PWAs and non-AIDS patients on the oncology ward and the general ward (Table 2). On the oncology ward, PWAs averaged over 1.5 nursing hours per day more than did other patients (6.7 vs 5.1 hours). On the general ward, the difference was about 0.7 hours, which, because of small numbers and wide variation among patients, was not significant. PWAs averaged less nursing time on the general ward than on the oncology ward, whereas non-AIDS patients received more nursing care on the general ward; these differences were not significant.

Discussion

It is generally thought that PWAs in nursing homes require more and costlier care than do non-AIDS patients. Judging by the data presented here on average nursing time per patient day, this argument appears to be true at a hospital-based SNF in California.

Our major finding is that PWAs received about 20% more nursing time than did other SNF patients. We found a difference of 1 hour per day (1½ hour on the oncology ward) in a hospital-based SNF, where all care is likely to be far heavier than that at freestanding SNFs, which constitute the majority of such facilities.[21]

Because not much is known about SNF care for PWAs, we cannot say whether our sample is similar to other groups of PWAs in nursing homes. The level of direct nursing time per patient day for non-AIDS patients at this SNF is double the level defining SNF care in California. Both PWAs and non-AIDS patients appear to approximate the designation of functionally dependent nursing home patients with complex needs.[16]

Despite expectations that hospital-based SNFs, which are paid by Medicaid at higher rates, deliver heavier care than do freestanding SNFs, data from a freestanding AIDS-dedicated SNF[1] suggest nursing times per patient day nearly as high as those we found at a hospital-based SNF. The difference between AIDS and non-AIDS care should be larger in freestanding SNFs, where the average patient is a chronically ill older woman receiving minimal nursing care, primarily from allied health personnel.[22] Thus, our comparison of care for PWAs and care for non-AIDS patients in the hospital-based SNF is a conservative test of the differences in nursing care for these two groups. Our small sample size limits the power of the analysis, however, and a more refined analysis of which patients and diagnoses account for care differences will require a study of a larger number of patients from a representative sample of facilities. Future research should focus on subsamples of PWAs in nursing homes, particularly intravenous drug users and ethnic minorities. □

Acknowledgments

The research reported herein was supported by the California Department of Health Services, Office of AIDS (contract no. 98-91877). The interpretations and conclusions are those of the authors and should not be attributed to the funding agency.

A version of this paper was presented at the 116th annual meeting of the American Public Health Association, Boston, Mass, November 1988.

We acknowledge the efforts of Robyn East, Thomas McKenzie, and Elaine Weston. We thank Dr Philip R. Lee for his guidance and support.

References

1. Swan JH, Benjamin AE. Nursing costs of skilled nursing facility care for AIDS. *AIDS Public Policy J.* 1990;5:64–67.
2. Rango N. Nursing home care for people with AIDS. In: Rogers DE, Ginzberg E, eds. *The AIDS Patient: An Action Agenda.* Boulder, Colo: Westview; 1988:35–41.
3. Edwards KS. Where will all the patients go? a statewide look at facilities for AIDS patients. *Ohio Med.* 1988;84:373–380.

4. Perry S, Maretta RF. AIDS dementia: a review of the literature. *Alzheimer Dis Assoc Disord.* 1987;1:221–235.
5. Wyatt A. Caring for persons with AIDS in geriatric nursing homes: geriatric nursing home care can work for persons with AIDS. *Health Soc Work.* 1988;13:156–157.
6. Benjamin AE. Long term care and AIDS: perspectives from experience with the elderly. *Milbank Q.* 1988;66:415–443.
7. Beresford L. Alternative outpatient settings of care for people with AIDS. *QRB.* 1989;15:9–16.
8. McNally L, Beck LM. A chronic care approach to health and social services for people with AIDS. *J Palliat Care.* 1988;4(4):96–99.
9. Crystal S. Persons with AIDS and older people: common long-term care concerns. In: Riley MW, Ory MG, Zablotsky D, eds. *AIDS in an Aging Society: What We Need to Know.* New York: Springer; 1989:147–166.
10. Intergovernmental Health Policy Program. *AIDS: A Public Health Challenge.* Washington, DC: George Washington University; 1987.
11. Bulkin W, Brown L, Fraioli D, Giannattasio E, McGuire G, Tyler P, Friedland G. Hospice care of the intravenous drug user AIDS patient in a skilled nursing facility. *J Acquired Immune Defic Syndromes.* 1988;1:375–380.
12. Benjamin AE, Swan JH. Nursing home care for persons with HIV illnesses. *Generations.* 1989;13(4):63–64.
13. May K. AIDS and the nursing home. *Nurs Homes.* 1987;36(5):8–11.
14. Kerschner P. Caring for persons with AIDS in geriatric nursing homes: weighing the costs of care. *Health Soc Work.* 1988; 13(2):157–158.
15. Adams HR. Financial problems inherent in the admission of AIDS patients into long term care facilities. *The J Leg Med.* 1989; 10(1):89–101.
16. Kearns JM. *The 1987 Evaluation and Update of the Staffing Criteria for the Criteria-Based Model.* Washington, DC: US Dept of Health and Human Services, Division of Nursing, Health Resources and Services Administration; 1987.
17. Gebhardt RB. Should nursing homes take AIDS patients? *Nurs Homes.* 1986; 35(2):13–15.
18. Meyer D. Costing nursing care with the GRASP system. In: Shaffer FA, ed. *Costing Out Nursing: Pricing Our Product.* New York, NY: National League for Nursing; 1985:55–67.
19. Shulman LC, Mantell JE. The AIDS crisis: a United States health care perspective. *Soc Sci Med.* 1988;26(10):979–988.
20. Centers for Disease Control. *HIV/AIDS Surveillance: U.S. AIDS Cases Reported Through October 1990.* Atlanta, Ga: Centers for Disease Control; November 1990.
21. Harrington C, Preston S, Grant LA, Swan JH. Trends in state nursing home bed capacity and occupancy in the 1978–1989 period. Presented at the 118th annual meeting of the American Public Health Association; October 1990; New York City.
22. Institute of Medicine, Committee to Study the Role of Allied Health Personnel. *Allied Health Services: Avoiding Crises.* Washington, DC: National Academy Press; 1989.

Risk Factors Associated with the Classification of Unspecified and/or Unexplained Causes of Death in an Occupational Cohort

Donna L. Cragle, PhD, and Amie Fetcher, DVM, MPH

ABSTRACT

In a matched case-control study of an occupational cohort in East Tennessee, data from the death certificates of 608 cases and controls were abstracted and analyzed in order to investigate possible risk factors associated with the certification on the death certificate of an ill-defined cause of death. There was a very strong association between the ill-defined classification and the certifier, especially if the certifier was a medical examiner (OR = 10.4, 95% CI: 6.0, 18.0). (*Am J Public Health.* 1992;82:455–457)

Introduction

Results from cohort studies of workers at Department of Energy (DOE) facilities in Oak Ridge, Tenn, indicated unusual increases in the rate of deaths due to unexplained and/or unspecified causes (*International Classification of Diseases,* 8th version, codes 780 to 796).[1,2] The standardized mortality ratio for deaths coded to these disease categories was 3.05 (338 observed vs 110.92 expected). This increase may have reflected actual increases in the incidence of these unexplained causes of death, variations in the quality and accessibility of health care, and/or inaccuracies recorded on death certificates for the underlying cause of death. These inaccuracies most commonly involve problems related to making the distinction between the underlying cause of death, the immediate cause of death, the manner of death, and conditions contributing to death.[3]

The objective of this study was to determine the association between various factors and a recorded ill-defined cause of death.

Methods

Study Population

The cases and controls were chosen from the pooled cohort of all White male DOE Oak Ridge facility workers who worked for more than 30 days between 1943 and 1982 and died between 1945 and 1982. The cohort of workers was com-

Donna L. Cragle is with Oak Ridge Associated Universities, Oak Ridge, Tenn. Amie Fetcher is with the University of Tennessee, Knoxville.

Requests for reprints should be sent to Donna L. Cragle, PhD, Director, Center for Epidemiologic Research, Oak Ridge Associated, PO Box 117, Oak Ridge, TN 37831-0117.

This paper was submitted to the Journal February 19, 1991, and accepted with revisions June 19, 1991.

TABLE 1—Distribution of Cases and Controls

	Cases	Controls
Year of birth, median (range)	1908 (1880–1942)	1908 (1880–1942)
Year of death, median (range)	1970 (1950–1982)	1970 (1948–1982)
Average age at death, y	62.0	62.0
Age range at death, y	24–91	25–92
Born in Tennessee,[a] %	60.7	50.7
Lived in Tennessee,[a] %	78.0	56.0
Died in Tennessee,[a] %	77.5	56.3

[a]Significant difference between cases and controls ($P < .001$).

TABLE 2—Odds Ratios for Significant Factors by Age Groupings

	Group								
	All Cases and Controls			<65 y			>65 y		
	Odds Ratio	95% Confidence Interval		Odds Ratio	95% Confidence Interval		Odds Ratio	95% Confidence Interval	
Died in Tennessee	3.3	1.9,	5.6	4.2	2.0,	8.9	3.3	1.2,	8.7
Died at home	2.5	1.6,	4.0	1.8	1.01,	3.3	3.4	1.4,	8.2
Single	1.7	1.1,	2.7	2.4	1.3,	4.3	0.7[a]	0.3,	1.7
No autopsy	3.1	1.5,	6.4	3.7	1.6,	8.6	3.2[a]	0.5,	22.1
Certifier									
Medical examiner	10.4	6.0,	18.0	8.6	4.2,	17.6	29.6	8.0,	110.0
Coroner	5.7	3.0,	11.0	7.7	3.2,	18.3	6.1	1.7,	21.5
Other	7.6	2.1,	27.0	5.4	1.3,	22.6		...[a]	

[a]Not significant.

posed of more than 70 000 White males, of whom approximately 18 000 were known to have died by the end of 1982.

Six hundred and eight "cases" (individuals who had died of unexplained and/or unspecified causes) were identified, and the same number of controls (individuals who had died of specific causes) were selected. Cases and controls were matched by date of birth (± 2 years), year of death (± 2 years), and facility of employment. Information abstracted from the death certificates of the cases and controls included marital status, state of birth, county/state of residence, county/state of death, place of residence (inside or outside of the city limits), job classification (professional, skilled, or unskilled), location of death (hospital, other institution, home, or other place), type of hospital (rural, urban, teaching, or Veterans Administration [VA]), hospital status (dead on arrival [DOA], emergency room [ER], or inpatient), autopsy examination (yes or no), and certifier (attending physician, medical examiner, coroner, or other).

Data Analysis

Descriptive information about the differences between the cases and controls was determined from frequency distributions for the classification variables. Matched conditional logistic regression analysis using the SAS-PHGLM[4] program was used to analyze associations, control for confounding, and determine odds ratios (ORs) and 95% confidence intervals (CIs) for the various classification variables.

Results

On the basis of the frequency distributions for year of birth, year of death, and age at death, the cases were matched well with the controls. These distributions are presented in Table 1. Data for cases and controls by state of birth, residence, and death were collapsed into two categories, (1) "Tennessee" and (2) "not Tennessee." All the frequency distributions for these single factors were significantly different between cases and controls ($P < .001$).

SAS-PHGLM was used to analyze a full model that included all variables with complete sets of data. The variables in this model were state of death, residence (inside or outside of city limits), place of death, marital status, job classification, length of employment, certifier, and autopsy (yes or no) (Table 2). The OR from this full model for having died in Tennessee and having had an unexplained and/or unspecified cause of death recorded on the death certificate was 3.3 (95% CI = 1.9, 5.6).

A marital status of single (never married, divorced, or widowed) accounted for 30.3% of the cases, compared with 21.5% of the controls. As a factor by itself, this variable was significant ($P < .001$), with an overall OR of 1.7 (95% CI = 1.1, 2.7).

Place of death was one of the most significant variables in the study. The overall OR for dying at home was 2.5 (95% CI = 1.6, 4.0). For those individuals dying in hospitals, the location of the hospital was analyzed with respect to state (Tennessee or not Tennessee), type of hospital (urban, rural, VA, or teaching), and hospital status (DOA, ER, or inpatient). Of all deaths in Tennessee, 59.9% occurred in hospitals; of all out-of-state deaths, 56.1% occurred in hospitals. The type of hospital was not significantly related to the cause of death recorded on the death certificate. However, the status of DOA was significant (OR = 2.7; 95% CI = 1.4, 5.1). Hospital status was unknown for 34.4% of all deaths because many death certificate forms did not request this information.

Autopsies were performed on 9.7% of all workers in the study, and the autopsy rate was significantly different between cases and controls (2.8% and 16.6%, respectively). Overall, the OR associated with no autopsy was 3.1 (95% CI = 1.5, 6.4). Even when an autopsy was performed, the information was not always used to complete the death certificate. Autopsy findings were recorded on the death certificates of only 18.8% of the cases that were autopsied (3/16), compared with 51.5% (51/99) of the controls.

Of all the variables considered, the individual completing the death certificate was the most significant. The overall OR for certification by a medical examiner was 10.4 (95% CI = 6.0, 18.0). Coroners and other certifiers had ORs of 5.7 (95% CI = 3.0, 11.0) and 7.6 (95% CI = 2.0, 27.0), respectively.

Because the cases and controls were matched by year of birth and year of death, age at death could not be considered as an independent factor in the full

model. Instead, the population was divided into two subgroups: (1) "young" (65 years old or less) and (2) "old" (over 65 years old). The full PHGLM model was then used for each of these subgroups. Certification by a medical examiner was a much stronger risk factor for the "old" group (OR = 29.6; 95% CI = 8.0, 110.0) than for the "young" group (OR = 8.6; 95% CI = 4.2, 17.6). The OR for having died at home was also higher for the "old" group (OR = 3.4; 95% CI = 1.4, 8.2) than for the "young" group (OR = 1.8; 95% CI = 1.0, 3.3). Single marital status at the time of death (OR = 2.4; 95% CI = 1.3, 4.3) and not having been autopsied (OR = 3.7; 95% CI = 1.6, 8.6) were both significant risk factors for the "young" group but not the "old" group (Table 1).

Discussion

One of the most surprising findings in this study was the very significant association between the certifier and the classification of cases. For all deaths, the OR for a medical examiner recording an ill-defined cause of death on the certificate was 10.4 (95% CI = 6.0, 18.0). Certification by a coroner or other certifier was also a very significant factor.

The OR of 3.1 (95% CI = 1.5, 6.4) that was associated with not having been autopsied was not surprising. One would have expected that the immediate and underlying causes of death would be determined from a postmortem examination. However, it was surprising that the autopsy findings were not recorded on all death certificates, even as a follow-up report of a pending investigation. The findings were not recorded in 15.0% of all autopsies.

When the study population was divided into "young" (65 years and under) and "old" (over 65 years) subgroups, the risk factors were qualitatively and quantitatively different for each group. The significant factors associated with having had an unexplained and/or unspecified cause of death on the death certificate for the "young" population were (1) death in Tennessee, (2) death at home, (3) single marital status, (4) no autopsy, and (5) a death certificate filled out by a medical examiner, coroner, or other certifier. For the "old" population, significant factors were (1) death in Tennessee, (2) death at home, and (3) certification by a medical examiner, coroner, or other certifier. Single marital status for those dying over 65 years of age was protective, although this was not significant. These findings were unexpected. One would have thought that older, single men might have died alone at home and not have received an autopsy. Under these circumstances, they would have been more likely to have an ill-defined cause of death recorded on the death certificate, especially in light of the high risk associated with certification by a medical examiner for this age group.

The results of this study indicate that the individual completing the death certificate, regardless of qualifications, can have a very important influence on the rate of reporting of certain diseases. In this case, the disease of interest was unexplained and/or unspecified illnesses. Place of death and hospital status were also significant factors associated with certification of an ill-defined cause of death. □

Acknowledgments

This report concerns work undertaken as part of the Health and Mortality Study of Department of Energy Workers being conducted by Oak Ridge Associated Universities with the collaboration of the School of Public Health, University of North Carolina at Chapel Hill under Contract No. DE-AC05-76OR00033 between the Department of Energy, Office of Energy Research, and Oak Ridge Associated Universities.

The authors acknowledge the vital statistics offices of the individual states as the source of death record data and appreciate the offices' technical support of this research. The authors are solely responsible for the data analyses and the interpretation of the results.

References

1. Frome EL, Cragle DL, McLain RW. Poisson regression analysis of the mortality among a cohort of World War II nuclear industry workers. *Radiat Res.* 1990; 123:138–152.
2. Wing S, Shy CM, Wood JL, et al. Mortality among workers at Oak Ridge National Laboratory: evidence of radiation effects in follow-up through 1984. *JAMA.* 1991;265: 1397–1402.
3. Kircher T, Anderson RE. Cause of death: proper completion of the death certificate. *JAMA.* 1987;258:349–352.
4. SAS Institute Inc. *SUGI Supplemental Library User's Guide, Version 5 Edition.* Cary, NC: SAS Institute Inc. 1986;437–466.

Performance of the Reflotron in Massachusetts' Model System for Blood Cholesterol Screening Program

Stephen Havas, MD, MPH, MS, Robert Bishop, MS, Lauren Koumjian, MSc, Joel Reisman, MS, and Susan Wozenski, MPH, JD

ABSTRACT

The precision, accuracy, and durability of the Reflotron were evaluated by the Massachusetts Model Systems for Blood Cholesterol Screening Program. Screenings were conducted in diverse community settings over 16 months. Fingerstick samples from 10 428 individuals were tested. None of the four analyzers met the 1992 standards for precision, although two met the 1992 standards for accuracy. More than 40% of Reflotron values differed from the reference laboratory values by upwards of 5%. More than 16% of individuals were misclassified in terms of their risk category. All four instruments malfunctioned during the project. (*Am J Public Health.* 1992;82:458–461)

Introduction

At least a dozen portable cholesterol-testing instruments have been designed for use in community and office settings. These instruments have made mass screenings for this risk factor feasible and thus are now in widespread use for this purpose.

To address questions concerning the reliability of this technology, the National Heart, Lung, and Blood Institute supported three projects entitled the Model Systems for Blood Cholesterol Screening Program (MSBCSP).[1] A primary MSBCSP objective was to test the accuracy, precision, and durability of portable analyzers under the usual operating conditions associated with screenings. This article reports findings from Massachusetts' MSBCSP concerning the Reflotron, the lengthiest field evaluation thus far reported.

Methods

From October 1987 to January 1989, the Massachusetts Department of Public Health conducted a blood cholesterol screening project in Chelsea, inner-city Boston, Worcester, and Lowell. The project screened 10 428 individuals in community settings using fingerstick samples analyzed by a Reflotron. Venipuncture blood samples were drawn on 972 participants for comparison.

Most screenings were conducted at work sites although other settings were also used. Two Reflotrons were transported to each site. Three levels of quality control (QC) material were analyzed at the beginning of each screening, after every 10th patient sample, and at the end. Values for QC materials were considered out of range based on Westgard's rules.[2] A reference laboratory at the University of Massachusetts at Lowell, which had met the Centers for Disease Control's Lipid Standardization Program requirements, analyzed the venipuncture serum samples.

Results

For Reflotrons 1, 3, and 4, the weighted average of all values met the current precision goal set by the National Cholesterol Education Program's Laboratory Standardization Panel (LSP) for all three QC levels, but it fell short of the 1992 LSP standards for all levels.[3] For Reflotron 2, the weighted average of the values met the current LSP standards for all QC levels, but the 1992 standards were attained for only two levels (see Table 1).

The bias of the four instruments varied from –3.3% to +3.2%. The differences among the Reflotrons were statistically significant (ANOVA $F = 25.4$, $P < .0001$). All four analyzers met the current LSP standards; only two met the 1992 standards.

As demonstrated in Figure 1, monthly bias on the Reflotron fluctuated considerably. For example, the bias ranged from –6.0% to +2.9% on Reflotron 1. Similar variations were seen on the other Reflotrons.

Figure 2 plots the differences in individual results between the Reflotron and the reference laboratory. One Reflotron 1 reading was 80% higher than the reference laboratory value, a 2.12 mmol/L difference. At the opposite extreme, another Reflotron 1 reading was 57% lower than the reference value, a difference of –3.41 mmol/L. The other plots demonstrate similar discrepancies. Overall, more than 40% of the Reflotron values differed from the

Stephen Havas, Robert Bishop, Lauren Koumjian, and Joel Reisman are all with the Massachusetts Department of Public Health. Susan Wozenski is with the University of Massachusetts at Lowell. Dr. Havas is also with the Department of Epidemiology and Preventive Medicine at the University of Maryland School of Medicine in Baltimore.

Requests for reprints should be sent to Dr. Stephen Havas, Department of Epidemiology and Preventive Medicine, University of Maryland School of Medicine, 655 W Baltimore Street, Baltimore, MD 21201.

This paper was submitted to the journal December 26, 1990, and accepted with revisions June 24, 1991.

venous values by upwards of 5%. The frequency of differences greater than 5% varied from 33% to 47% among the Reflotrons.

The Reflotrons incorrectly classified from 9.2% to 18.5% of individuals regarding their blood cholesterol category. For three Reflotrons, misclassification tended to result in false reassurance; for the other Reflotron, misclassification resulted in false warnings. The differences among the Reflotrons were statistically significant ($\chi^2 = 22.0$, $P < .005$). Overall, 16.3% of individuals were misclassified by the four Reflotrons.

An unanticipated finding was the malfunctioning of the analyzers. Reflotron 2 had to be returned to the manufacturer when its QC results consistently fell out of range. Subsequently, Reflotron 3 had to be returned for similar reasons. On the last two scheduled days for the project, neither Reflotron 1 nor Reflotron 4 analyzed the QC materials within acceptable ranges, resulting in cancellation of the screenings.

TABLE 1—Precision of the Reflotrons, by Percent

Reflotron	Solomon Park High CV (n)	Solomon Park Medium CV (n)	Solomon Park Low CV (n)
1	3.5 (243)	3.1 (232)	3.5 (218)
2	2.6 (116)	2.6 (84)	3.5 (87)
3	4.4 (24)	3.8 (23)	4.7 (22)
4	4.4 (77)	3.5 (61)	3.7 (51)

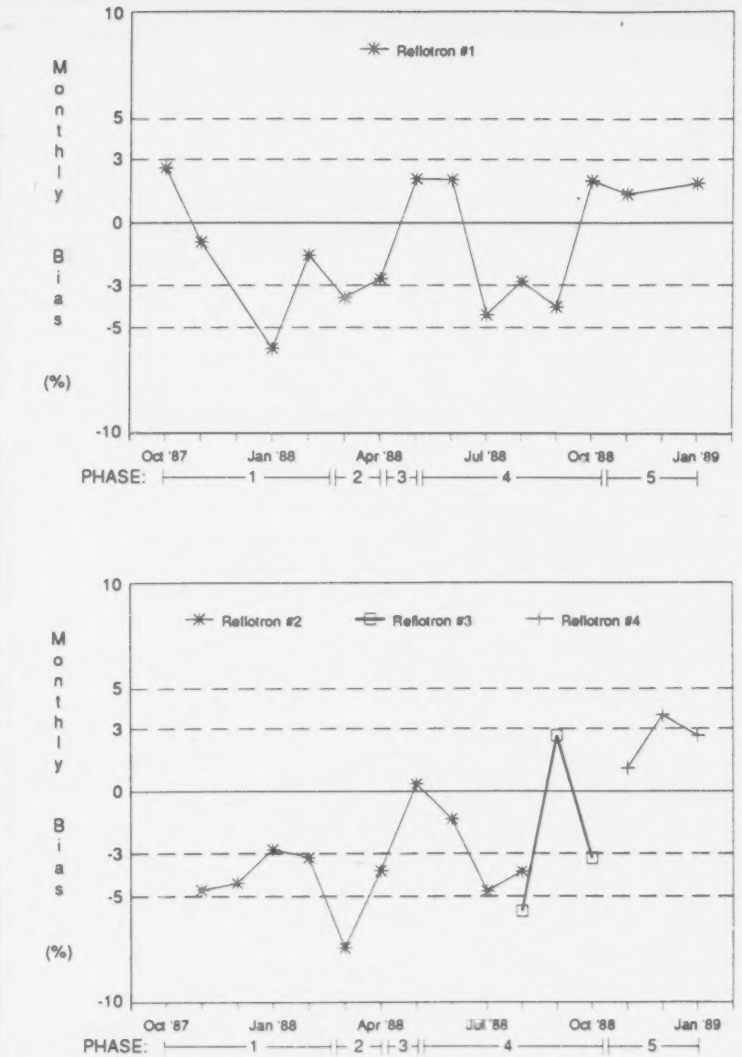


FIGURE 1—Monthly bias plots for the four Reflotrons.

Discussion

The Reflotron has been marketed aggressively for use in community screening programs. The marketing has focused heavily on the instrument's relatively low cost, ease of operation, and accuracy. This strategy has resulted in the widespread use of this instrument in blood cholesterol screenings.

As shown in Table 2, the Reflotron has been studied previously using various settings, sample sizes, and methodologies.[4–21] When placed in the context of these other studies, our results lead to concerns about the widespread usage of the Reflotron. The first concern is that the instrument is not sufficiently precise. Although it does meet current standards, it clearly will not be able to meet the 1992 standards without substantial modifications.

Our second concern is that the Reflotron does not consistently produce accurate results. Accuracy was found to vary greatly over time, both among and within Reflotrons. Our study is the first to show that monthly bias can fluctuate from positive to negative and vice versa for the same Reflotron. This finding is important because it means one cannot simply apply a "correction" factor to screening values based on overall bias.

A disturbing finding is the large individual discrepancies and the frequency of error, which make it difficult to ascertain blood cholesterol levels correctly and to classify individuals appropriately. These inaccuracies occurred despite all four Reflotrons meeting the current LSP standards for precision and accuracy.

Based on our findings and on similar data in two other recent studies,[20,21] it is likely that screening programs will misclassify many people due to the current level of inaccuracy of the Reflotron. This could result in various adverse effects. For example, those with high or borderline high blood cholesterol levels who were falsely reassured may not have the follow-up that current treatment guidelines recommend.[22] On the other hand, those

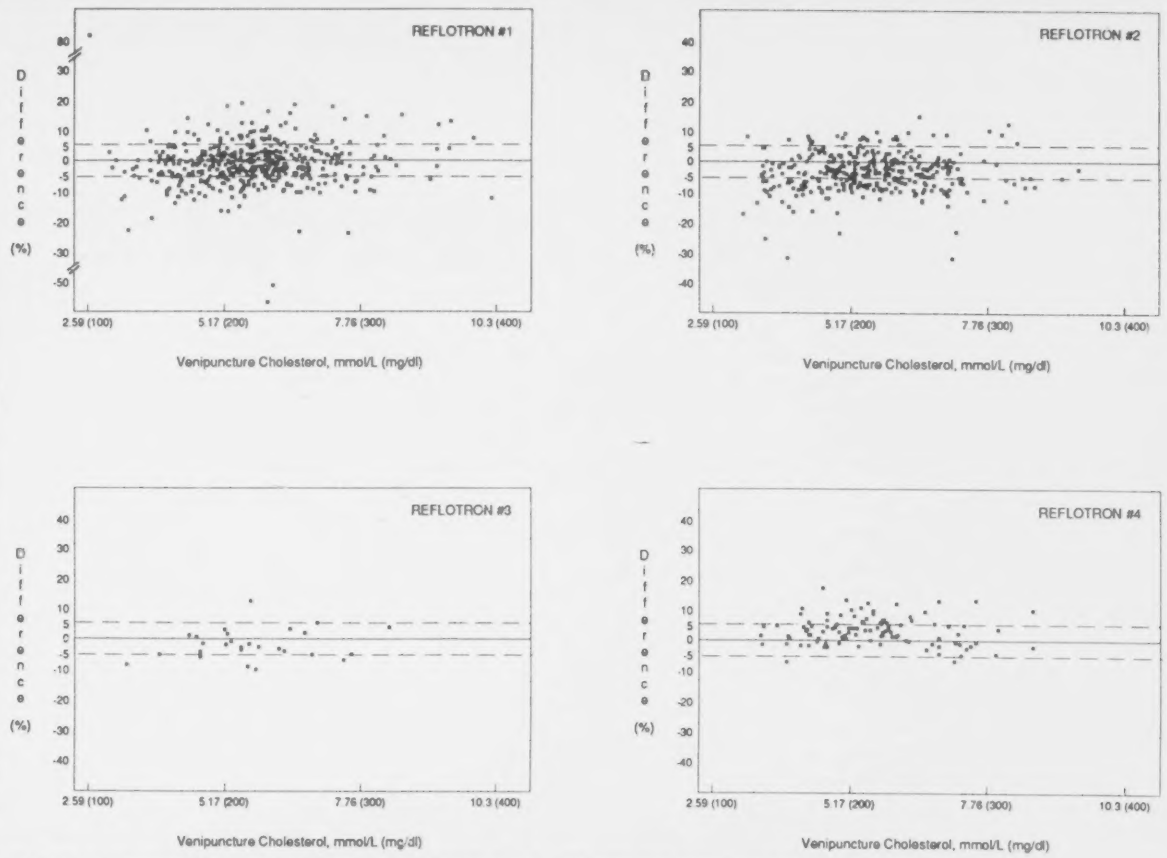


FIGURE 2—Percent difference between results from the four Reflotrons and reference laboratory.

TABLE 2—Previous Studies of the Reflotron

Investigator	Date	Sample Size (n)	Coefficient of Variation (%)	Setting	Correlation Coefficient (r)	Bias (%)
Koch[4]	1987	109	2.3–3.8	lab	0.95	"significant negative bias"
Von Schenck[5]	1987	256	5.4–9.4	lab/M.D. office	0.92–0.95	"acceptable"
Rohac[6]	1987	98	2.5–2.9	lab	0.96	+0.5%
Sanguini[7]	1988	29	not reported	lab	0.98	"small negative proportional bias"
Sedor[8]	1988	204	3.4–4.0	lab	0.95	"usually within 5%"
James[9]	1988	30	not reported	field	0.92	–3.0
Phillips[10]	1988	80	1.1–1.8	lab	0.97	+4.1
Burke[11]	1988	83	2.0–2.9	lab	0.97	+4.5
Boerma[12]	1988	86	1.6–3.4	lab	0.96	–2.2
Price[13]	1988	98	1.8–4.8	lab	0.92	not identified
Seftel[14]	1988	117	0.7–1.0	lab	0.98	+6.0
Rastam[15]	1988	304	3.1–4.7	lab/field	0.92–0.93	+4.1 to +4.9
Pearson[16]	1988	345	9.1–9.9	field	not reported	–1.1
Bachorick[17]	1989	1298	2.0–4.3	lab	0.92–0.96	–0.8 to –7.8 (pooled data for 8 Reflotrons)
Lefebvre[18]	1989	486	"within 5.0"	field	not reported	–1.8 to –10.5
Assmann[19]	1989	85	1.5–2.5	lab	not reported	–0.2
Bachorick[20]	1990	290	4.0–7.1	field	0.89	+6.0 (pooled data for 50 Reflotrons)
Kaufman[21]	1990	96	13.3	lab	not reported	+5.2 to +6.3

who were falsely warned may visit their physicians unnecessarily, incurring additional tests, costs, and anxiety.

The frequent breakdown of the Reflotrons is also worrisome. For analyzers to be useful to screening programs, they must be precise, accurate, and durable. Our Reflotrons were not sufficiently durable to be considered optimal.

The public's enthusiasm for blood cholesterol screening programs is likely to continue for quite some time. Such programs have the potential to improve awareness and control rates for high blood cholesterol substantially. To provide the quality of blood cholesterol measurements that this important public health problem warrants, however, requires either improvements in the Reflotron or the use of other, more reliable analyzers. □

Acknowledgments

This study was funded by the National Heart, Lung, and Blood Institute grant number RO1HL37832. The authors wish to acknowledge Gene Rogers, PhD (University of Massachusetts at Lowell), for his thoughtful review of this manuscript and the helpful suggestions he provided. We also express our appreciation to Cheryl Blik, MS (former director, Division of Health Promotion Sciences, Massachusetts Department of Public Health), and Lily Hsu, MS (Model Systems Project Director), for their administrative oversight of this project and Robert Nicolosi, PhD (University of Massachusetts at Lowell) for performing the reference laboratory analyses.

References

1. Havas S, Greenland P, Wones R. Addressing unanswered questions about population cholesterol screenings: the Model Systems for Blood Cholesterol Screening Program. *Am J Prev Med.* 1989;5:337–346.
2. Westgard JO, Klee GG. Quality assurance. In: Tietz NW, ed. *Textbook of Clinical Chemistry.* Philadelphia, Pa: WB Saunders; 1986:424–458.
3. National Cholesterol Education Program. *Recommendations for Improving Cholesterol Measurement: A Report from the Laboratory Standardization Panel of the National Cholesterol Education Program.* Washington, DC: US Govt Printing Office; 1990. US Public Health Service publication PHS 90–2964.
4. Koch TR, Mehta U, Lee H, et al. Bias and precision of cholesterol analysis by physician's office analyzers. *Clin Chem.* 1987;33:2262–2267.
5. Von Schenck H, Treichl L, Tilling B, et al. Laboratory and field evaluation of three desktop instruments for assay of cholesterol and triglyceride. *Clin Chem.* 1987;33:1230–1232.
6. Rohac M, Gabl F. Comparison of two solid-phase chemistry systems: Reflotron and Ektachem DT 60. *J Clin Chem Clin Biochem.* 1987;25:811–821.
7. Sanguini LC, Ceriotti F, Franzini C, et al. Determination of serum cholesterol on the Reflotron (BBR) dry chemistry analyser. *Ann Biol Clin.* 1988;46:312–321.
8. Sedor FA, Holleman CM, Heyden S, et al. Reflotron cholesterol measurement evaluated as a screening technique. *Clin Chem.* 1988;34:2542–2545.
9. James R, Tyler C, Henrikson D. An evaluation of the accuracy of the Reflotron system in the field. *Med J Aust.* 1988;149:130–131.
10. Phillips S, Wyndham L, Shaw J, et al. How accurately does the Reflotron dry-chemistry system measure plasma total cholesterol levels when used as a community-screening device? *Med J Aust.* 1988; 149:122–125.
11. Burke JJ, Fischer PM. A clinician's guide to the office measurement of cholesterol. *JAMA.* 1988;259:3444–3448.
12. Boerma GJM, Gelderland J, van Gorp I, et al. Use of the Reflotron system for cholesterol assay in capillary blood, venous blood, and serum: evaluation of accuracy and lot-to-lot reagent comparability. *Clin Chem.* 1988;34:2117–2119.
13. Price CP, Koller PU. A multicentre study of the new Reflotron system for the measurement of urea, glucose, triacylglycerols, cholesterol, gamma-glutamyltransferase and haemoglobin. *J Clin Chem Clin Biochem.* 1988;26:233–250.
14. Seftel HC, Panz VR, Baker SG, et al. Determination of cholesterol and triglycerides in blood: a comparison between wet chemistry methods and a dry chemistry analyser. *Ann Clin Biochem.* 1988;25:176–180.
15. Rastam L, Admire JB, Frantz ID Jr, et al. Measurement of blood cholesterol with the Reflotron analyzer evaluated. *Clin Chem.* 1988;34:426.
16. Pearson R, Dusenbury LJ, Bakes-Martin R, et al. Evaluation of a simple method for measuring blood cholesterol levels using non-laboratory observers. *Am J Med.* 1988;85:369–374.
17. Bachorik PS, Bradford RH, Cole R, et al. Accuracy and precision of analyses for total cholesterol as measured with the Reflotron cholesterol method. *Clin Chem.* 1989;35:1734–1739.
18. Lefebvre RC, Lasater TH, McKinlay SM, et al. Performance characteristics of a blood cholesterol measuring instrument used in screening programs. *Public Health Rep.* 1989;104:266–270.
19. Assmann G, Brinkers H, Schulte H, et al. Comparison of the reflectance method (Reflotron reflectance photometer) with the absorbance method (automatic analysers) for the determination of cholesterol. *J Clin Chem Clin Biochem.* 1989;27:961–966.
20. Bachorik PS, Rock R, Cloey T, et al. Cholesterol screening: comparative evaluation of on-site and laboratory-based measurements. *Clin Chem.* 1990;36:255–260.
21. Kaufman HW, McNamara JR, Anderson KM, et al. How reliably can compact chemistry analyzers measure lipids? *JAMA.* 1990;263:1245–1249.
22. Report of the National Cholesterol Education Program's Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. *Arch Intern Med.* 1988;148:36–69.

American Journal of Public Health

86789
0008
CONGRESSIONAL INFORMATION SVC
4520 EAST-WEST HIGHWAY
BETHESDA MD 20814
ED KIANKA

April 1992, Vol. 82, No. 4 Established 1911

Editorials

521 Serving All *M. Susser*

522 Report Card on Our National Response to the AIDS Epidemic—Some A's, Too Many D's *D. E. Rogers*

Commentary

525 Standards for Peer Evaluation: The Hospital Quality Assurance Committee *S. E. Feldman and D. W. Roblin*

Featuring HIV/AIDS

528 Factors Associated with AIDS Risk Behaviors among High School Students in an AIDS Epicenter *H. J. Walter, R. D. Vaughan, M. M. Gladis, D. F. Ragin, S. Kasen, and A. T. Cohall*

533 AIDS Education for Drug Abusers: Evaluation of Short-term Effectiveness *J. McCusker, A. M. Stoddard, J. G. Zapka, C. S. Morrison, M. Zorn, and B. F. Lewis*

541 HIV Infection in Intravenous Drug Users Entering Drug Treatment, United States, 1988 to 1989 *D. M. Allen, I. M. Onorato, T. A. Green, and the Field Services Branch of the Centers for Disease Control*

547 HIV Prevalence and Risk Behavior among Injecting Drug Users Who Participate in "Low-Threshold" Methadone Programs in Amsterdam *C. Hartgers, A. (J.A.R.) van den Hoek, P. Krijnen, and R. A. Coutinho*

561 Impact of HIV Infection on Mortality and Accuracy of AIDS Reporting on Death Certificates *N. A. Hessol, S. P. Buchbinder, D. Colbert, S. Scheer, R. Underwood, J. L. Barnhart, P. M. O'Malley, L. S. Doll, and A. R. Lifson*

564 Seroprevalence of HIV-Type 1 in a Northern California Health Plan Population: An Unlinked Survey *R. A. Hiatt, F. J. Capell, and M. S. Ascher*

567 The Validity of Self-reported HIV Antibody Test Results *J. McCusker, A. M. Stoddard, and E. McCarthy*

569 Cognitive and Attitudinal Impacts of a University AIDS Course: Interdisciplinary Education as a Public Health Intervention *R. P. Strauss, I. B. Corless, J. W. Luckey, C. M. van der Horst, and B. H. Dennis*

573 Ineffectiveness of AIDS Education and HIV Antibody Testing in Reducing High-Risk Behaviors among Injection Drug Users *D. A. Calsyn, A. J. Saxon, G. Freeman, Jr, and S. Whittaker*

576 The Inpatient AIDS Unit: A Preliminary Empirical Investigation of Access, Economic, and Outcome Issues *M. C. Fahs, G. Fulop, J. Strain, H. S. Sacks, C. Muller, P. D. Cleary, J. Schmeidler, and B. Turner*

578 Duration of Medicaid AIDS Hospitalizations—Variation by Season, Stage, and Year *L. E. Markson, B. J. Turner, and T. R. Fanning*

581 Decline in Human Immunodeficiency Virus Seropositivity and Seroconversion in US Navy Enlisted Personnel: 1986 to 1989 *F. C. Garland, E. D. Gorham, S. O. Cunnion, M. R. Miller, L. L. Balazs, and the Navy HIV Working Group*

584 Fear of Dying and HIV Infection vs Hepatitis B Infection *L. J. Schneiderman and R. M. Kaplan*

(continued page 507)

APHA
American Public Health Association

Featuring HIV/AIDS

Contents

American Journal of Public Health

"Information for Authors" appeared on page 134 of the January 1992 issue and may also be obtained from the Journal Editorial Office. A shortened and revised version is printed in this issue.

American Journal of Public Health
1015 Fifteenth Street, NW
Washington, DC 20005

APHA
American Public Health Association

April 1992, Vol. 82, No. 4 ISSN: 0090-0036 Established 1911

587 What Is the Dentist's Occupational Risk of Becoming Infected with Hepatitis B or the Human Immunodeficiency Virus? *E. I. Capilouto, M. C. Weinstein, D. Hemenway, and D. Cotton*

590 Prevalence of HIV Antibodies in Transsexual and Female Prostitutes *B. Modan, R. Goldschmidt, E. Rubinstein, A. Vonsover, M. Zinn, R. Golan, A. Chetrit, and T. Gottlieb-Stematzky*

592 Predictors of Condom Accessibility among Hispanics in San Francisco *B. V. Marin and G. Marin*

595 Legal Needle Buying in St. Louis *W. M. Compton III, L. B. Cottler, S. H. Decker, D. Mager, and R. Stringfellow*

Articles

552 The Potential Role of Custody Facilities in Controlling Sexually Transmitted Diseases *D. Cohen, R. Scribner, J. Clark, and D. Cory*

557 Occupational Injuries among Minors Doing Farm Work in Washington State: 1986 to 1989 *N. J. Heyer, G. Franklin, F. P. Rivara, P. Parker, and J. A. Haug*

Public Health Briefs

597 Risk Factors Associated with Antibodies to Leptospires in Inner-city Residents of Baltimore: A Protective Role for Cats *J. E. Childs, B. S. Schwartz, T. G. Ksiazek, R. R. Graham, J. W. LeDuc, and G. E. Glass*

600 Characteristics Related to Elderly Persons' Not Eating for 1 or More Days: Implications for Meal Programs *E. A. Frongillo, Jr, B. S. Rauschenbach, D. A. Roe, and D. F. Williamson*

603 The South Karelia Air Pollution Study: Acute Health Effects of Malodorous Sulfur Air Pollutants Released by a Pulp Mill *T. Haahtela, O. Marttila, V. Vilkka, P. Jäppinen, and J. J. K. Jaakkola*

606 The Epidemiology of Low Back Pain in an Adolescent Population *T. L. Olsen, R. L. Anderson, S. R. Dearwater, A. M. Kriska, J. A. Cauley, D. J. Aaron, and R. E. LaPorte*

Notes from the Field

609 Informed Consent and Hysterectomy: Enhancing the Right to Know *B. K. Rose*

610 "A Star Is Born" in Ypsilanti *D. Share, C. Givens, D. Davies, E. Eley, and J. Chesler*

611 Health Corners: Reducing Chronic Disease Risks among Black Public Housing Residents in the Nation's Capital *M. L. Rivo and K. Gray*

Other Journal Departments

508 Masthead

509 What *AJPH* Authors Should Know

510 Advertisers' Index

527 Erratum

613 Letters to the Editor

617 Association News: Leadership Roster

632 Book Corner

639 Job Opportunities

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, **published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.**

APHA OFFICERS

Joyce C. Lashof, MD, *President*

C. William Keck, MD, MPH, *Immediate Past President*

Helen Rodriguez-Trias, MD, *President-Elect*

Arnold E. Greenberg, SM, *Vice-President (USA)*

Trevor Hancock, MB, MHSc, *Vice-President (Canada)*

Julio Teja Perez, MD, *Vice-President (Latin America)*

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Frances M. Veverka, RS, MPH, *Chair, Executive Board*

George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*

William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD

Frances M. Veverka, RS, MPH (1992), *Chair*

Marilyn Aguirre-Molina, EdD, MS (1993)

Brian A. Dobrow, MPH (1992)

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Kristine M. Gebbie, BSN, MN (1994), *Vice-Chair*

Alan R. Hinman, MD, MPH (1995)

Janet Heinrich, DrPH, MPH (1993)

Margaret H. Jordan, RN, MPH (1992)

C. William Keck, MD, MPH, *Immediate Past President*

Margo D. Kerrigan, MPH (1995)

Joyce C. Lashof, MD, *President*

Barry S. Levy, MD, MPH (1993)

Helen Rodriguez-Trias, MD, *President-Elect*

Allan G. Rosenfield, MD (1994)

Reed V. Tuckson, MD (1994)

Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO

Willam H. McBeath, MD, MPH, *Executive Director*

Joanne E. Lukomnik, MD, *Chair, Action Board*

George Pickett, MD, MPH, *Chair, Program Development Board*

George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc., 1015 Fifteenth St., NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50,000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health,* 1015 Fifteenth St., NW, Washington, DC 20005.

Information on Manuscript Preparation is published in this Journal in January. The Journal publishes reports of original research, demonstrations, evaluations, and other articles covering current aspects of public health. Manuscripts are accepted for consideration with the understanding that they, or their essential substance, have not been published previously or accepted for publication and are not under consideration elsewhere.

Subscription price, individual, is $80/year US, $120/year international; **institutional,** $160/year US, $240/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to American Journal of Public Health, 1015 Fifteenth St., NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact: Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact: Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols. 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS Bibliographic Retrieval Service Inc, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc., Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Factors Associated with AIDS Risk Behaviors among High School Students in an AIDS Epicenter

ABSTRACT

Background. A greater understanding of the determinants of risky behaviors is an essential precursor to the development of successful AIDS prevention programs for adolescents.

Methods. A survey measuring AIDS-related behaviors, beliefs, and knowledge was administered to a sample of 531 10th-grade students residing in an AIDS epicenter.

Results. Of the 56.8% of students reporting past-year involvement in sexual intercourse, 67.3% reported unprotected intercourse with low-risk partners, 1.3% reported unprotected intercourse with high-risk partners, and 6.6% reported a past-year history of a sexually transmitted disease. Students whose friends had intercourse and never or inconsistently used condoms, who personally sanctioned intercourse involvement, who believed that the majority of their peers had intercourse, and who perceived low preventive action self-efficacy, were 5.1, 3.0, 2.1, 3.7, and 2.8 times more likely, respectively, to score in the riskier categories of an AIDS behavior index.

Conclusions. These findings suggest that addressing socioenvironmental influences on risky and preventive behaviors may prove to be the most effective AIDS prevention strategy among adolescents. (*Am J Public Health.* 1992;82:528–532)

Heather J. Walter, MD, MPH, Roger D. Vaughan, MS, Madeline M. Gladis, PhD, Deborah Fish Ragin, PhD, Stephanie Kasen, PhD, and Alwyn T. Cohall, MD

Introduction

Adolescents, particularly those residing in or near one of the nation's acquired immunodeficiency syndrome (AIDS) epicenters, are at increasing risk of acquiring infection with human immunodeficiency virus (HIV), the virus that causes AIDS.[1] Although the total number of reported AIDS cases among 13- to 19-year-olds thus far is small, the caseload in this age group has increased by 40 percent in the past year alone.[1] Moreover, one fifth of all AIDS cases has been reported among 20- to 29-year-olds,[1] many of whom acquired HIV infection as teenagers. Accordingly, the development, implementation, and evaluation of AIDS prevention programs for inner-city adolescents has become a recent public health priority.[2]

A greater understanding of the determinants of risk-related and preventive behaviors in target populations is an important precursor to the development of successful AIDS prevention programs.[3] In the past several decades, a number of theoretical models have been advanced to predict involvement in a variety of preventive actions.[4–6] Recently, constructs derived from several of these models have been applied theoretically and empirically to risky and preventive behaviors specific to AIDS,[7–11] and therefore these models were selected for inclusion in this investigation.

The health belief model[4] posits that individuals who believe that they are personally susceptible to the target disease, that the consequences of having the disease would be severe, and that the benefits of engaging in preventive actions outweigh the barriers are more likely to adopt preventive behaviors than are individuals not holding these beliefs. Social cognitive theory[6] suggests that one's appraisal of one's self-competence (i.e., self-efficacy) pertaining to preventive behaviors generates the self-confidence necessary for the successful performance of these behaviors. Finally, a recently proposed model of reference group–based social influence[11] hypothesizes that behavioral norms (beliefs about how people typically act) and values (beliefs about how people ought to act) interact to generate a normative standard to which individuals aspire to conform, thereby influencing the adoption and maintenance of preventive behaviors.

To investigate the associations between beliefs derived from the seven constructs outlined above (susceptibility, severity, benefits, barriers, self-efficacy, norms, and values) and involvement in AIDS risk and preventive behaviors, a survey was administered in the spring of 1990 to a sample of 10th-grade students residing in an AIDS epicenter. The purpose of the survey was to identify correlates of risky and preventive behaviors that could be targeted for modification or reinforcement, respectively, in preven-

Heather J. Walter, Stephanie Kasen, and Alwyn T. Cohall are with the College of Physicians and Surgeons, Columbia University, New York, NY. Roger D. Vaughan, Madeline M. Gladis, and Deborah Fish Ragin are with the New York State Psychiatric Institute, New York, NY.

Requests for reprints should be sent to Heather J. Walter, MD, MPH, New York State Psychiatric Institute, Box 29, 722 W 168th St, New York, NY 10032.

This paper was submitted to the Journal April 12, 1991, and accepted with revisions November 12, 1991.

tion programs. The survey analyses had four goals: first, to ascertain the prevalence of AIDS risk and preventive behaviors; second, to determine whether these behaviors varied by age, gender, race/ethnicity, school attended, or AIDS-related knowledge; third, to describe the cross-sectional associations of AIDS risk and preventive behaviors with beliefs derived from the constructs identified above; and fourth, to determine whether the observed belief–behavior associations varied according to the demographic and knowledge variables identified above.

Methods

Subjects

Tenth-grade students in 2 of the 14 public academic high schools in a New York city borough were selected as the eligible study population. The two schools were selected on the basis of their willingness to participate in the survey and their racial/ethnic similarity to the entire population of the 14 schools. Passive parental consent was required for students' participation in the survey; the nonconsent rate was 5.5% in one school and 6.5% in the other. The mean age of participants (n = 531) was 16.0 years (SD, 1.0; range, 13 to 21 years); 44.3% of the participants were male. The racial/ethnic distribution of the participants was 59.2% Black, 28.3% Hispanic, and 12.5% other (primarily non-Hispanic Whites); the participants were predominantly from working-class or welfare-recipient families. The demographic distribution of the survey participants was similar to that of the eligible study population.

Measurements

Survey forms were distributed by research staff to all consenting, present 10th-graders enrolled in required general education courses in the spring semester of 1990. Survey completion required one class period; responses to the survey were anonymous.

AIDS-risk behaviors. The students were asked to complete five items regarding past-year involvement in AIDS risk and preventive behaviors, including use of illicit intravenous drugs, involvement in sexual intercourse, number and risk status of sexual partners, and consistency of condom use during sexual intercourse. A high-risk sexual partner was defined as one who was believed by the respondent to have a history of male homosexual intercourse or illicit intravenous drug use. An AIDS behavior index, ranging from lower to higher risk of acquiring infection with HIV, was created from these items according to the epidemiological model of HIV transmission proposed by Hearst and Hulley.[12] Index scores ranged from 0 to 6; the definition of each risk level is presented in Table 1.

Beliefs derived from the health belief model. The students were asked to complete two items pertaining to perceived susceptibility to AIDS and three items pertaining to perceived severity of AIDS; and 15 and 17 items, respectively, constituting additive scales of the perceived benefits of and barriers to engaging in AIDS-preventive behaviors (i.e., sexual abstinence, consistent condom use, sexual monogamy, and knowing the sexual and drug use histories of sex partners). The internal consistency (alpha coefficient) of the benefits scale was 0.81, and that of the barriers scale, 0.77.

Beliefs derived from social cognitive theory. The students were asked to complete 13 items constituting an additive scale of perceived self-efficacy pertaining to the preventive actions identified above. The internal consistency (alpha coefficient) of the scale was 0.80.

Beliefs derived from the model of social influence. The students were asked to complete eight items assessing their beliefs about how many of their four closest friends and what proportion of the students in their grade were involved in the past year in the preventive behaviors identified above (i.e., beliefs about norms). They were also asked to complete four items assessing whether they personally believed that people their age should be involved in these behaviors (i.e., beliefs about values).

AIDS-related knowledge. The students were asked to complete 12 items pertaining to the transmission and prevention of HIV infection. These items were selected on the basis of their measurement characteristics from an original 72-item scale; the internal consistency (alpha coefficient) of the original scale was 0.82.

Statistical Analysis

The analytic strategy comprised four steps. First, descriptive statistics were calculated to provide prevalence data for the investigated behaviors and beliefs, and bivariate relations between the behaviors and beliefs were assessed by computing odds ratios (ORs) and associated 95% confidence intervals (CIs). Second, multiple logistic regression was employed to assess differences in the distribution of AIDS behavior index scores across the five demographic and knowledge variables. Logistic rather than linear regression was employed because of the multinomial nature of the distribution of the AIDS behavior index scores. Accordingly, the AIDS behavior index scores were dichotomized to reflect lower versus higher risk. Third, multiple logistic regression again was used to assess the unique associations between beliefs derived from the seven investigated theoretical constructs and the AIDS behavior index scores. Finally, interaction variables, created by multiplying each demographic term or terms by the significantly associated beliefs, were added to the regression model to assess interaction effects.

TABLE 1—Proportions (in percent) of Students in each Category of the AIDS Behavior Index (n = 531)

Category[a]	Males	Females
0	38.4	46.7
1	8.9	10.6
2	8.2	1.6
3	13.7	30.1
4	26.0	6.5
5	0.0	0.0
6	4.8	4.5

[a]Category 0 = sexual abstinence and no illicit intravenous (IV) drug use.
Category 1 = sexual intercourse consistently using a condom with one low-risk partner and no illicit IV drug use.
Category 2 = sexual intercourse consistently using a condom with two or more low-risk partners and no illicit IV drug use.
Category 3 = sexual intercourse with one low-risk partner without or inconsistently using a condom and no illicit IV drug use.
Category 4 = sexual intercourse with two or more low-risk partners without or inconsistently using a condom and no illicit IV drug use.
Category 5 = sexual intercourse consistently using a condom with one or more high-risk partners and no illicit IV drug use.
Category 6 = sexual intercourse with one or more high-risk partners without or inconsistently using a condom, or illicit IV drug use, or history of a sexually transmitted disease.

Results

AIDS-Risk and Preventive Behaviors

Nearly two thirds (64.6%: 76.7% of males and 55.6% of females) of this sample of 10th-grade students reported ever having had sexual intercourse. The modal age of initiation of intercourse was 11 years for males and 15 years for females. Three fifths (59.3%: 59.5% of males and 59.2% of

TABLE 2—Multivariate Odds Ratios (ORs)[a] and 95% Confidence Intervals (CIs) for Past-Year Higher AIDS Risk,[b] Given Modeled Beliefs (n = 531)

Modeled Beliefs	OR (95% CI)
Friends had intercourse in past year (Two to four vs none or one)	5.1 (2.5,10.7)
Friends never or inconsistently used condoms in past year (Two to four vs none or one)	3.0 (1.8,4.9)
People my age should have intercourse (Should vs should not)[c]	2.1 (1.3,3.5)
Peers had intercourse in past year (≥ 1/2 vs < 1/2)	3.7 (1.1,12.3)
Preventive action self-efficacy (Bottom vs top tertile)	2.8 (2.1,3.9)

[a]Odds ratios are adjusted for age, gender, race/ethnicity, school, and AIDS-related knowledge.
[b]Higher AIDS risk includes categories 3–6 (see Table 1).
[c]Dichotomized responses to a 7-point Likert scale (should = 1 to 3; should not = 4 to 7)

TABLE 3—Proportions of Students (in percent) Endorsing "Not at all sure," "A little sure," and "Somewhat sure" Responses to Self-Efficacy Items (n = 531)

Self-Efficacy Items	Males	Females
Refuse intercourse with longtime date	65.6	67.5
Refuse intercourse with partner you want to fall in love with you	67.6	60.4
Refuse intercourse if "high" or "drunk"	67.5	46.8
Refuse intercourse if partner's sex and drug histories unknown	64.2	38.8
Refuse intercourse if condom not used	69.2	31.1
Purchase condoms	36.9	54.4
Use condoms when "high" or "drunk"	56.2	38.0
Use condoms correctly	44.1	31.1
Use condoms consistently	43.2	29.5
Remain sexually monogamous	61.8	8.0
Ask sex partner about drug use history	40.0	22.7
Ask that condom be used during intercourse	36.1	16.0
Ask sex partner about sex history	37.5	14.9

females) reported ever having used (or having had their partner use) a condom during intercourse.

Table 1 presents the proportions of students reporting past-year involvement in each category of AIDS-related behaviors. No students reported use of illicit intravenous drugs; less than half reported sexual abstinence. Only one fifth of sexually active students reported protected intercourse with one low-risk partner.

Belief–Behavior Correlations

Significant bivariate associations were observed between beliefs derived from all seven investigated theoretical constructs (i.e., susceptibility, severity, benefits, barriers, self-efficacy, norms, and values) and the AIDS risk behavior scores. However, in a multivariate analysis in which beliefs derived from all seven constructs were entered into the model along with age, gender, race/ethnicity, school attended, and AIDS-related knowledge, beliefs derived from only three (i.e., norms, values, and self-efficacy) of the seven constructs were significantly associated with involvement in higher-risk behaviors (Table 2). In the presence of beliefs derived from these three constructs, beliefs derived from the four health belief model constructs (i.e., perceived susceptibility to and severity of AIDS, and perceived benefits of and barriers to engaging in preventive actions) were not significantly associated with higher-risk behaviors.

The analyses revealed a significant main effect of age on AIDS behavior index scores: older students were slightly more likely to be in the higher categories of risk than were younger students (OR = 1.2 per year; 95% CI, 1.1 to 2.1). The analyses failed to provide any evidence of significant main effects of gender, race/ethnicity, school attended, or AIDS-related knowledge on the index scores, nor was there evidence of interaction between any of these variables and the belief–behavior associations.

The great majority of the students believed that half or more of their peers and friends had had sexual intercourse in the past year; half of the students believed that half or more of their friends had never or inconsistently used condoms in the past year. Three fifths of the males, but only one quarter of the females, believed that people their age should have intercourse. Substantial proportions of the students (in general, more males than females) were not sure that they could perform certain behaviors essential for effective AIDS prevention (Table 3).

Discussion

Our principal finding is that a substantial proportion of a sample of 10th-grade students residing in an AIDS epicenter reported involvement in behaviors that place them at risk for acquiring infection with HIV. Beliefs derived from three theoretical constructs (i.e., norms, values, and self-efficacy) were cross-sectionally associated with riskier behaviors. In the presence of beliefs derived from these three constructs, beliefs derived from health belief model constructs were not significantly associated with higher AIDS risk.

Before we consider the implications of these results, some methodological issues should be addressed. First, because of the self-report nature of the survey, the veracity of students' responses may be questioned. However, responses to the survey were anonymous to maximize accurate self-disclosure. In addition, there was high agreement (above 98%) of responses to similarly worded items assessing the same behavioral domains that were dispersed throughout the survey. The validity of responses also is strengthened by the similarity of the behavior prevalence rates reported in this study to those reported in other studies.[13–18] Finally, there is evidence suggesting that questionnaire data are as accurate as interview data in assessing sexual-behavior prevalence rates among groups of subjects.[19]

Second, although the high survey participation rates obtained in this study suggest that the findings may be generalizable to similar populations of students, the findings may not be applicable to students attending schools unwilling to participate in such surveys, to students with a different demographic profile, or to teen-

agers who are habitually absent or have withdrawn from school.

A number of recent studies[13–18] conducted among various populations of adolescents have suggested that substantial proportions of teenagers and young adults engage in behaviors that increase their risk for acquiring infection with HIV. The most common behavioral profile among sexually active adolescents in this sample was a teenager who had unprotected sexual intercourse with partners believed to be at low risk for carrying HIV. Because in the absence of HIV antibody testing one cannot be certain that one's partner indeed is uninfected,[20] and because seropositivity rates in AIDS epicenters are estimated to be high,[21] this behavioral profile (which is applicable to over two thirds of the sexually active students in this sample) conveys nontrivial risk to these teenagers. Of particular concern because of their greatly increased risk of HIV infection[12] are the 1.3% of the sexually active students in this sample who reported unprotected intercourse with high-risk partners, and the 6.6% who reported a history of sexually transmitted disease (similar percentages also have been reported elsewhere).[13–18]

Although a number of studies have documented the prevalence of AIDS risk behaviors among adolescents, fewer have investigated theoretically or empirically suggested determinants of these behaviors. The findings of two studies recently conducted among teenagers have been informative in this regard. Hingson et al.[8] showed that beliefs about the benefits and barriers of condom use and beliefs about susceptibility to getting AIDS were associated cross-sectionally with condom use, as is predicted by the health belief model. Kasen et al.[10] demonstrated that intercourse refusal and condom use self-efficacy were correlated with involvement in sexual intercourse and condom use among adolescents, net of the influence of beliefs derived from health belief model constructs, as is predicted by social cognitive theory.

The findings from this study demonstrate that although bivariate associations between beliefs derived from health belief model constructs and AIDS risk behaviors were observed, these associations were attenuated in the presence of certain socioenvironmental factors (i.e., beliefs about norms, values, and self-efficacy). The results are similar to those reported from a number of studies examining factors influencing drug use among teenagers,[22–24] and suggest that individual/cognitive interventions (e.g., those based upon the health belief model) emphasizing the disease-related knowledge and beliefs of individuals as they exist apart from the social world may have only small effects on teenagers' motivations to initiate preventive actions. Instead, a more productive strategy may be to address salient socioenvironmental influences on preventive behaviors. For example, the data from this study suggest that teenagers could be helped to clarify their personal values relating to involvement in preventive actions, and an understanding of the effects of outside influences (e.g., peer pressure) on those values could be facilitated. Adolescents' misconceptions about the prevalence of risky and preventive behaviors among their peers could be corrected. Preventive action self-efficacy could be enhanced by empowering youths with the skills necessary to successfully perform risk-reducing behaviors. Skills in which teenagers may be particularly deficient include questioning sexual partners about sex and drug-use histories, using condoms correctly and consistently, and refusing offers to engage in sexual intercourse in certain high-pressure social situations.

In summary, teenagers (particularly those residing in or near AIDS endemic areas) are at risk for acquiring infection with HIV. Because certain socioenvironmental factors (i.e., beliefs about norms, values, and self-efficacy) may strongly influence involvement in AIDS risk behaviors, further elucidation of the social influence process among adolescents should become a prevention research priority. For the present, it may prove to be beneficial for practitioners to reinforce values incompatible with involvement in AIDS risk behaviors, correct misconceptions about the normative status of such behaviors, and teach skills enabling effective, sustained preventive action. □

Acknowledgments

This research was supported by NIMH/NIDA grant 5-P50-MH43520 to Heather J. Walter, MD, MPH, Principal Investigator, AIDS Prevention for Adolescents in School, and Anke A. Ehrhardt, PhD, Principal Investigator, HIV Center for Clinical and Behavioral Studies.

The authors wish to thank Richard Neugebauer, PhD, and Judith Rabkin, PhD, for reviewing the manuscript.

References

1. *HIV/AIDS Surveillance Report.* Atlanta, Ga: Centers for Disease Control; December 1990.
2. Flora JA, Thoresen CE. Reducing the risk of AIDS in adolescents. *Am Psychol.* 1988;43:965–970.
3. Kirscht JP. Preventive health behavior: a review of research and issues. *Health Psychol.* 1983;2:277–301.
4. Becker M. The health belief model and personal health behavior. *Health Educ Monogr.* 1974;2(special issue).
5. Ajzen I, Fishbein M. *Understanding Attitudes and Predicting Social Behavior.* Englewood Cliffs, NJ: Prentice-Hall; 1980.
6. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice-Hall; 1986.
7. Becker M, Joseph J. AIDS and behavioral change to avoid risk: a review. *Am J Public Health.* 1988;78:384–410.
8. Hingson RW, Strunin L, Berlin BM, Heeren T. Beliefs about AIDS, use of alcohol and drugs, and unprotected sex among Massachusetts adolescents. *Am J Public Health.* 1990;80:295–299.
9. Bandura A. Perceived self-efficacy in the exercise of control over AIDS infection. Presented at the Conference on Women and AIDS: Promoting Health Behaviors; National Institutes of Mental Health and Drug Abuse research conference; 1987; Bethesda, Md.
10. Kasen SK, Vaughan RD, Walter HJ. Self efficacy for AIDS preventive behaviors among tenth grade students. *Health Educ Q.* In press.
11. Fisher JD. Possible effects of reference group–based social influence on AIDS-risk behavior and AIDS prevention. *Am Psychol.* 1988;43:914–920.
12. Hearst N, Hulley SB. Preventing the heterosexual spread of AIDS: are we giving our patients the best advice? *JAMA.* 1988;259:2428–2432.
13. Goodman E, Cohall AT. Acquired immunodeficiency syndrome and adolescents: knowledge, attitudes, beliefs, and behaviors in a New York City adolescent minority population. *Pediatrics.* 1989;84:36–42.
14. Anderson JE, Kann L, Holtzman D, Arday S, Truman B, Kolbe L. HIV/AIDS knowledge and sexual behavior among high school students. *Fam Plann Perspec.* 1990;22:252–255.
15. Hingson R, Strunin L, Berlin B. Acquired immunodeficiency syndrome transmission: changes in knowledge and behaviors among teenagers, Massachusetts statewide surveys, 1986 to 1988. *Pediatrics.* 1990;85:24–29.
16. Keller SE, Bartlett JA, Schleifer SJ, Johnson RL, Pinner E, Delaney B. HIV-relevant sexual behavior among a healthy inner-city heterosexual adolescent population in an endemic area of HIV. *J Adolesc Health Care.* 1991;12:44–48.
17. MacDonald NE, Wells GA, Fisher WA, et al. High-risk STD/HIV behavior among college students. *JAMA.* 1990;263:3155–3159.
18. Kotloff KL, Tacket CO, Clemens JD, et al. Assessment of the prevalence and risk factors for human immunodeficiency virus type 1 (HIV-1) infection among college students using three survey methods. *Am J Epidemiol.* 1991;133:2–8.
19. DeLamater J, MacCorquodale P. The ef-

fects of interview schedule variations on reported sexual behavior. *Sociol Meth Res.* 1975;4:215–236.
20. Cochran SD, Mays VM. Sex, lies and HIV. *N Engl J Med.* 1990;322:774. Letter.
21. Drucker E, Vermund ST. Estimating population prevalence of human immunodeficiency virus infection in urban areas with high rates of intravenous drug use: a model of the Bronx in 1988. *Am J Epidemiol.* 1989;130:133–142.
22. Jessor JR, Chase JA, Donovan JE. Psychosocial correlates of marijuana use and problem drinking in a national sample of adolescents. *Am J Public Health.* 1980;70:604–613.
23. Kandel DB. Epidemiological and psychosocial perspectives on adolescent drug use. *J Am Acad Child Psychiatry.* 1982;21:328–347.
24. Robinson TN, Killen JD, Taylor CB, et al. Perspectives on adolescent substance use. A defined population study. *JAMA.* 1987;258:2072–2076.

Thirty-First Hanford Symposium on Health and Environment—Call for Abstracts

Date: October 20 to 23, 1992
Location: Richland, Washington, USA
Title: The Laboratory and Epidemiology: The Development and Application of Biomarkers to the Study of Human Health Effects
Sponsors: US Department of Energy
Battelle, Pacific Northwest Laboratories
Department of Environmental Health, University of Washington
School of Public Health and Community Medicine

Contact: Dr. Lowell E. Sever
Symposium Chairman
Battelle Seattle Research Center
206/528-3348
Address for Abstracts: Dr. Ray W. Baalman
Life Sciences Center, MS K1-50
Battelle, Pacific Northwest Laboratories
Richland, WA 99352, USA
509/375-3665
Deadline: May 1, 1992

AIDS Education for Drug Abusers: Evaluation of Short-term Effectiveness

Jane McCusker, MD, DrPH, Anne M. Stoddard, ScD, Jane G. Zapka, ScD, Charles S. Morrison, PhD, Martha Zorn, MS, and Benjamin F. Lewis, EdD

ABSTRACT

Background. Interventions are needed to assist drug abusers in reducing risky drug and sexual behavior.

Methods. A randomized controlled trial compared three small-group AIDS educational interventions among 567 clients of a 21-day inpatient drug detoxification program: a two-session informational intervention, given either during the first (early) or second (late) week of treatment; and a six-session enhanced intervention. Changes in knowledge, attitudes, and psychomotor skills were assessed before and after each intervention, and behavioral outcomes were assessed at follow-up 10 to 18 weeks after admission.

Results. Immediately after the interventions, enhanced group members reported significantly greater self-efficacy to talk themselves out of AIDS-risky behavior; other knowledge and attitude scales did not differ by intervention. At follow-up, significant reductions in risky drug use were reported by all groups. Enhanced group members reported significantly greater reduction in injection frequency than did late informational subjects.

Conclusions. No beneficial effect was detected of delaying AIDS education for clients entering detoxification. At this early stage of follow-up, there is only weak evidence that an enhanced intervention improved outcomes. (*Am J Public Health.* 1992;82:533–540)

Introduction

Drug users, particularly intravenous drug users, are a critical target group for acquired immunodeficiency syndrome (AIDS) educational interventions.[1] In this study we focused on drug users admitted to a short-term, inpatient drug detoxification and rehabilitation program in central Massachusetts, with the expectation that this temporarily captive population might be more receptive to educational programs than one concurrently using drugs. Because the program is short, usually not more than 21 days, and most clients return to drug use following discharge, this seemed an important opportunity for an intervention that might have a "ripple effect" on clients' peers in the community in addition to the anticipated risk reduction among clients who participated in the program.

The objectives of the study were to compare the effects of two AIDS risk reduction interventions, informational and enhanced. The *informational* strategy was based on "standard" AIDS education programs commonly provided in treatment settings, which primarily impart information. In two 1-hour group sessions in which a specially designed video, lectures, homework, and class discussion were used, the subjects were provided with basic information about human immunodeficiency virus (HIV) transmission, prevention, symptoms, and testing. The subjects watched demonstrations of condom use and cleaning of drug paraphernalia with a real needle and syringe, but did not practice themselves.

Although the theoretical basis for the informational intervention included concepts from the Health Belief Model[2] and the theory of reasoned action,[3] the *enhanced* intervention emphasized this material more heavily. The enhanced intervention was based primarily on social cognitive theory[4] and relapse prevention theory.[5,6] A broader interpretation of relapse was used, which included not only relapse to drug use, but relapse to HIV-risky drug-use and sexual behaviors.

The enhanced intervention consisted of six 1-hour group sessions, followed by a 30-minute individual health education consultation, and focused on personal susceptibility, situational analysis, and skill building. Although the essential information about HIV and AIDS (prevention, transmission, etc.) was provided in the same manner as in the informational interventions, the enhanced classes focused on putting the knowledge into practice as well as on a more thorough discussion and practice of situations and skills. A variety of methods were used, including in-depth class discussions, additional homework, tension release exercises, role-playing, trigger tapes, peer feedback, and hands-on needle cleaning and condom exercises. The experiential learning was intended to enhance the subjects' self-efficacy regarding their ability to use skills that would assist them in initiating and maintaining AIDS harm reduction behaviors.

The study was designed to test hypotheses regarding the relative effectiveness of the informational versus the enhanced intervention. Because the two

The authors are with the University of Massachusetts AIDS Research Unit, University of Massachusetts School of Public Health, Amherst, Mass.

Requests for reprints should be sent to Jane McCusker, MD, DrPH, School of Public Health, Arnold 406, University of Massachusetts, Amherst, MA 01003.

This paper was submitted to the Journal April 29, 1991, and accepted with revisions November 27, 1991.

TABLE 2—Selected Characteristics of Cohort at Baseline: HIV Status and Risk Period Behaviors

	No.	%
HIV Status (n = 567)		
Never tested	252	45
Tested—positive	45	8
Tested—negative	241	43
Tested—result unknown	16	3
Risk Period Behaviors (n= 497)[a]		
Drugs used		
Heroin	402	81
Cocaine	400	81
Marijuana	282	57
Other opiates	178	36
Methadone	137	28
Composite drug-use risk behavior		
No injection	105	21
Injection, no borrowing	108	22
Borrowing, bleach always used	74	15
Borrowing, bleach used sometimes	119	24
Borrowing, bleach never used	84	17
(Injectors only)		
Frequency of injection		
Daily	351	90
At least weekly	21	5
Less than weekly	17	4
Number of sharing partners		
1	110	39
2–3	96	34
4–9	48	17
10+	25	9
Opposite-sex partners		
0	81	16
1	266	54
2–9	118	24
10+	31	6
Same-sex partners		
0	481	97
1+	15	3
Composite sexual risk behavior[b]		
No vaginal/anal contact	89	18
1 Partner, condom used always/sometimes	69	14
1 Partner, condom never used	188	38
2+ Partners	149	30

[a]Totals may be less than 497 due to missing data.
[b]Only penile-vaginal and penile-anal intercourse are considered.

interventions contained similar information, we hypothesized that they would be equally effective in achieving improved knowledge, but that the enhanced intervention would be more effective in changing attitudes, beliefs, skills, and behavior.

A second set of hypotheses was related to the timing of the informational intervention. We felt that the clients being detoxified on methadone might be less receptive to the intervention during the first week of detoxification, owing to the effect of higher doses of methadone. We therefore hypothesized that the informational intervention given early in the detoxification program (days 5–8) would be less effective, and the same intervention given at a later stage (days 10–15) would be more effective, because most clients would be at lower doses of methadone (5 or 10 mg/day) or completely drug-free. Because the enhanced program required a longer period of time, it was not feasible to compare an early and a late time of administration of the enhanced intervention. We therefore limited the intervention groups to three: early informational, late informational, and enhanced.

Methods

Baseline Characteristics of Study Sample

The majority of the subjects were male (67%), white (81%), high school graduates (70%), and gave English as their first language (92%). Of those whose first language was not English, 84% gave Spanish as their first language. The majority (80%) had had previous drug abuse treatment episodes. Prior HIV testing results and risk period behaviors are shown in Table 2.

The following baseline characteristics of subjects did not differ significantly at the $P<.01$ significance level by intervention group: gender; prior HIV antibody testing and result; levels of the summary drug-use and sexual risk behaviors (the latter stratified by gender); knowledge, attitudes and beliefs, and psychomotor skills.

Design

A blocked randomization design with pretest and multiple posttests was used. The interventions were allocated by date of admission, with all clients admitted during a specified period given the same intervention, regardless of their participation in the research. Each intervention admission period was planned to be 6 1/2 weeks long and the interventions were randomly ordered in blocks of three. Thus, each intervention would be given four times over the course of the study. In this way any seasonal fluctuation in admissions or secular trends in knowledge or behavior would be spread over all interventions. Furthermore, this length of admission period was intended to reduce the contamination between interventions by minimizing the number of transition periods. Small adjustments were made in the randomization schedule to balance the numbers of study subjects in the three study groups.

Table 1 shows the schedule of research contacts with subjects. Clients were approached for consent to participate on the first or second day of the program. The timing of the interviews during the treatment program was designed to assess subjects at baseline (T0), after the early informational but before the late informational intervention (T1), and after completion of both the late informational and enhanced interventions (T2). Follow-up interviews (T3) took place between 10 to 18 weeks after admission. Longer-term follow-up was carried out but is not reported here.

A nonintervention control group was not included. The purpose of this study was to evaluate the incremental effects of the enhanced over the more standard informational intervention, inasmuch as we believe that most drug treatment programs offer some basic form of AIDS education to their clients. In addition, we considered it unethical not to provide at least the in-

formational intervention to all subjects, in view of the severity of HIV disease and the highly susceptible client population. However, because the late informational group received no intervention before the T1 interview, they provided a control for the evaluation of the immediate effects of the early informational intervention.

The subjects received counseling regarding HIV antibody testing during their first week, and were given the option of HIV antibody testing at any subsequent point during their stay. If testing occurred, the results were released, with posttest counseling, after completion of the T2 interview, to avoid contamination of the intervention effect.

Outcomes

Primary study outcomes included selected attitudes and beliefs, knowledge, psychomotor skills, and drug-use and sexual behaviors. Selected secondary outcomes included program retention, HIV antibody testing, and initiation of longer-term drug abuse treatment following exit.

Four knowledge scales assessed knowledge of methods whereby HIV is transmitted, methods for prevention of transmission, ways to prevent progression to AIDS, and understanding of the HIV antibody test. Attitude and belief items or scales assessed perceived susceptibility to AIDS, benefits of risk reduction, attitudes toward slips in drug-use and sexual risk behaviors, and self-efficacy for eight relevant risk reduction behaviors. Psychomotor skill scales assessed the client's demonstration of correct methods for cleaning injection equipment with bleach and using a condom. (See Appendix for a description of variables).

The reporting period for drug-use and sexual behaviors at baseline and follow-up was the "risk period," defined as the most recent free-living period of at least 1 week and not more than 3 months.[7] Drug-use behaviors reported included type of drugs used and route of administration. Drug injectors were asked about two forms of sharing of injection equipment: borrowing of injection equipment previously used by others, and lending of injection equipment to others after first using it. Specific information requested included frequency of injection, borrowing, and lending of injection equipment, and cleaning of borrowed injection equipment with bleach before use; and numbers of persons who had previously used borrowed injection equipment ("sharing partners"). A composite drug-use risk variable was constructed from reported injection, borrowing, and bleach use. Sexual behaviors reported included numbers of sexual partners of the same and opposite sex and use of condoms during vaginal or anal sex; a composite sexual risk variable was constructed from these responses (Table 2).

TABLE 1—Schedule and Content of Research Contacts with Study Subjects

Time from Admission	Nature of Contact
1–2 days	Study introduction and informed consent
2–5 days	Baseline (T0) interview; baseline knowledge, attitudes, skills
7–10 days	Time 1 (T1) interview: repeat knowledge and attitudes: baseline behavior
11–15 days	Time 2 (T2) interview: repeat knowledge, attitudes, skills
10–18 weeks	Follow-up (T3) interview: follow-up behavior, attitudes, skills

Note. T1 interview took place after the early informational and before the late informational intervention, the T2 interview took place after all interventions; the T3 interview took place after the subject's discharge from the program.

Statistical Methods

The primary hypothesis-testing analyses followed the "intention to treat" principle, and included all subjects allocated to a given intervention, whether or not they actually received it as planned. Each analysis was computed using subjects who completed the relevant pre- and posttests (T0–T2) within 1 day of the protocol windows.

Each analysis compared the change in the outcome from pretest to posttest across study groups. For continuous and interval outcome measures, the difference in outcome between two time points was computed and study-group mean differences were compared by Student's t test or the one-way analysis of variance.[8] For ordinal categorical measures, changes in outcome were compared by a method based on the cumulative logit.[9]

Owing to the large number of hypothesis tests involved, statistical significance was determined by a P value of less than .01. All analyses were computed with the personal computer version of SAS.[10]

Enrollment and Attrition

All clients admitted to the detoxification program from July 25, 1988, to February 24, 1990, were eligible for study enrollment, unless they had enrolled or had received any of the study interventions during a prior admission.

During the recruitment period there were 961 admissions to the detoxification program. Of these admissions, 116 (12%) occurred during periods of nonenrollment due to understaffing, either of interviewers or the health educator, and 178 (18%) were readmissions of persons previously admitted during either the study period or the pretest of the study protocol, and who were therefore ineligible for study enrollment. Out of the 667 first admissions during enrollment periods, 4 clients were unable to complete the interviews because of language barriers, 13 left the program before their participation in the study could be requested, and 32 left against staff advice or were discharged before beginning the study protocol. Of the remaining 618 persons, 51 (8.3%) refused to participate, leaving 567 subjects in the study sample.

The following numbers of subjects completed scheduled in-house interviews within 1 day of the protocol window period: T0—557 (98%); T1—479 (84%); T2—415 (73%); and 335 (59%) completed the T3 follow-up interview within the scheduled window period of 18 weeks from admission. Behavioral outcomes were assessed among those 301 subjects who had completed the T1 interview within 1 day of the protocol window period and who also completed a T3 interview. Completers and noncompleters of any given interview did not differ by age, gender, race, education, intervention group, detoxification drug, or baseline level of the outcome variables. Only language was an important predictor of completion. Subjects whose first language was not English had significantly lower interview completion rates.

Results

Participation in Intervention, HIV Testing, and Further Treatment

The length of stay in the detoxification program varied from 1 to 30 days, with a median of 18 days, and was not

TABLE 3—Mean Levels (and Standard Deviation) of Knowledge, Attitudes, and Skills at T0 and T2 Interviews

	Scale Range	T0	T2	No.
Knowledge				
Transmission methods	0–10	5.11 (1.96)	6.97 (2.15)	409
Transmission prevention	0–8	5.14 (1.17)	5.52 (1.29)	409
Prevention of progression	0–6	1.60 (1.18)	2.24 (1.01)	409
HIV test knowledge	0–6	2.97 (1.06)	3.74 (0.98)	409
Attitudes and beliefs				
Susceptibility	2–8	5.00 (1.55)	4.67 (1.62)	386
Benefits	4–16	2.45 (1.67)	13.09 (1.69)	376
Sex slips recovery	1–4	2.97 (0.63)	3.06 (0.70)	406
Bleach slips recovery	1–4	2.37 (0.73)	2.65 (0.82)	406
Self-efficacy to:				
Avoid drugs	1–4	3.09 (0.74)	3.26 (0.62)	407
Refuse to lend works	1–4	3.45 (0.77)	3.57 (0.66)	330
Clean with bleach	1–4	3.33 (0.77)	3.61 (0.59)	331
Carry bleach	1–4	3.10 (0.93)	3.25 (0.79)	331
Avoid risky situations	1–4	3.31 (0.71)	3.33 (0.63)	399
Carry condoms	1–4	2.98 (1.09)	3.04 (0.99)	398
Use condoms	1–4	2.69 (1.09)	2.89 (1.00)	399
Self-talk	1–4	3.38 (0.69)	3.48 (0.64)	399
Psychomotor skills				
Bleach cleaning	0–3	1.29 (1.06)	2.38 (0.92)	306
Condom use	0–1	0.32 (0.47)	0.88 (0.32)	359

Note. T0 was the baseline interview; the T2 interview took place after all interventions.

associated with intervention. Most subjects (78%) were detoxified with methadone.

Altogether, 411 subjects (72%) attended all group sessions; complete attendance was lower for the enhanced group (68%) than for either informational group (74% and 75%), owing to the larger number of sessions in the enhanced intervention.

Overall, 48% of the subjects elected to be tested for HIV antibody while in treatment; 10% of these tested positive. There was a weak association of intervention with testing: 48% and 54% of the early and late informational groups, respectively, were tested, compared with 42% of the enhanced group ($P = .06$).

Of the 331 subjects who gave information on further treatment following exit from the program, 98 (30%) entered residential drug-free programs and 44 (13%) entered other drug abuse treatment programs; these proportions did not differ by intervention group.

Hypothesis Testing

We tested hypotheses related to intervention type and timing, with the latter hypotheses restricted to comparisons of the early versus late informational interventions among subjects detoxified on methadone. To simplify the presentation of results, we have not stratified by detoxification drug. These analyses (available on request) did not change our overall conclusions.

Immediate effects of interventions. At baseline, the subjects displayed moderate levels of knowledge, perceived susceptibility, and skills, but relatively high levels of other attitudes, particularly self-efficacy (Table 3). At the T1 interview, among subjects detoxified on methadone, the early informational group showed greater changes in 12 of the 16 knowledge and attitude items than did the late informational group; however, only knowledge of HIV transmission was statistically significant (data not shown). Members of both groups increased significantly in their knowledge of the test for HIV antibody and in their belief in their ability to recover from a failure to use bleach consistently. The subjects in the early informational group decreased significantly in their perceived susceptibility to AIDS and increased significantly in their knowledge of HIV transmission, knowledge of prevention of HIV infection, belief in the benefits of behavior change, self-efficacy to clean needles with bleach, and self-efficacy to use condoms.

By the T2 interview, the subjects in all intervention groups had increased significantly on almost all knowledge scales, but the differences between groups were not statistically significant (Table 4). There were no significant between-group differences in improvement in attitudes except in self-efficacy to talk oneself out of risky behavior (self-talk). On this item, the subjects in the enhanced group improved more than did subjects in the other two groups. All three groups showed increases in their belief in the benefits of behavior change, in their ability to recover from failure to use bleach consistently, in their self-efficacy to avoid using drugs, and in both their needle-cleaning and condom-use skills. Perceived susceptibility decreased in all three groups, but the change was statistically significant only for the early informational and enhanced groups.

Behavioral effects of interventions. At follow-up, there was a substantial reduction in risky drug-use behavior (Table 5). Only 44% of those interviewed reported drug injection, compared with 77% at baseline. A higher percentage of participants also reported abstinence from vaginal or anal sexual contact at follow-up (33%) than at baseline (19%).

The analysis based on cumulative logits gives an odds ratio for a "safer" behavior category at follow-up relative to baseline for each intervention group (Table 6). For example, subjects in the early informational group were 4.7 times more likely to report a lower frequency of injection at follow-up than at baseline. For most of the behaviors, all groups reported safer behaviors at follow-up than at baseline (odds ratios all greater than unity), with relatively greater behavior changes for frequency of injection and the composite drug-use variable. Less marked changes were found for frequency of borrowing and bleaching and for sexual behavior, and some of the confidence intervals included unity. However, except for frequency of injection, there were no significant differences in changes in behavior between the intervention groups. The odds ratio for the enhanced group is significantly greater than that for the late informational group, but not significantly greater than that for the early informational group.

Secondary Analyses

We repeated the analyses described above, limiting the sample to those subjects who had completed the intervention sessions to which they had been allocated as well as completing the T0 and T2 interviews within 1 day of the protocol windows (n = 339). We obtained essentially the same results as in the primary analyses by "intention to treat."

Because of the weak effect of the intervention on the decision to be HIV-antibody tested, we repeated the analyses of behavior change, using the composite drug-use and sexual behavior variables, with stratification by decision to be tested. Again, the results concerning association between intervention and behavior change remained unchanged, although the odds ratios were closer to unity. In all groups, those who were tested made slightly greater changes in their risky drug behavior than did those not tested, although the difference was not statistically significant. There remained no significant change in risky sexual behavior for any intervention when we stratified by testing.

Discussion

The results of this study show substantial improvement in AIDS knowledge and skills and reductions in risky behavior in the study subjects (achieved primarily through high rates of abstinence from injection and increased celibacy). The comparison of the early and late informational groups at T1 was a controlled evaluation of the informational intervention, and demonstrated changes in knowledge and attitudes attributable to the intervention, although only knowledge of HIV transmission achieved statistical significance. Thus, we conclude that the informational intervention was moderately effective.

Our hypotheses regarding the timing of the informational intervention were not confirmed; the early informational intervention appeared to be at least as effective as the late informational intervention. The "drugged" state of subjects during the earlier stages of methadone detoxification apparently did not compromise their ability to learn from the intervention, as shown by their significant gains in knowledge from T0 to T1. Clinical observations suggested that the period immediately following withdrawal from methadone, which usually occurred during the second week of the program and coincided with the late informational intervention, was particularly stressful for many subjects and may have compromised the effectiveness of this intervention.

The relative effectiveness of the enhanced intervention versus the informational interventions was disappointing. Immediately after the interventions, enhanced-group members reported significantly greater self-efficacy to talk oneself out of AIDS-risky behavior; other attitude items and scales did not differ by group. At follow-up, the enhanced group had made greater reductions in the frequency of injection, but other measures of risky drug-use behavior, including safer injection practices among those who continued to inject, showed no differential effect by intervention. Possibly, injection frequency is reported more reliably than needle sharing and cleaning behavior. Longer-term follow-up is needed to assess outcomes more completely; preliminary analyses of follow-up after 6 months indicate a stronger effect of the enhanced intervention in reducing relapse to injection.[11]

A number of factors may have contributed to a dilution of the treatment effect. Drug users who enter drug abuse treatment are a subpopulation that differs by factors such as education and race from other subpopulations of drug users.[7] Their motivation to change their behavior may have overwhelmed the potential to achieve differences between the interven-

TABLE 4—Mean (and Standard Deviation) Changes in Knowledge, Attitudes, and Skills from T0 to T2, by Intervention Group

	Early Informational (n = 123)[a]	Late Informational (n = 134)[b]	Enhanced (n = 152)[c]
Knowledge			
Transmission methods	1.33* (2.70)	1.94* (2.62)	2.19* (2.60)
Transmission prevention	0.30 (1.67)	0.42* (1.49)	0.40* (1.60)
Prevention of progression	0.73* (1.29)	0.68* (1.26)	0.55* (1.19)
HIV test knowledge	0.63* (1.31)	0.98* (1.18)	0.71* (1.24)
Attitudes and beliefs			
Susceptibility	−0.34* (1.15)	−0.28 (1.34)	−0.37* (1.17)
Benefits	0.60* (1.84)	0.60* (1.61)	0.70* (1.78)
Sex slips recovery	0.16 (0.75)	0.04 (0.68)	0.08 (0.67)
Bleach slips recovery	0.21* (0.76)	0.22* (0.89)	0.40* (0.82)
Self-efficacy to			
Avoid drugs	0.19* (0.77)	0.16* (0.69)	0.17* (0.74)
Refuse to lend works	0.07 (0.91)	0.16 (0.83)	0.12 (0.68)
Clean with bleach	0.15 (0.85)	0.31* (0.72)	0.36* (0.70)
Carry bleach	0.23* (0.78)	0.11 (0.83)	0.11 (0.86)
Avoid risky situations	−0.03 (0.78)	0.03 (0.69)	0.04 (0.75)
Carry condoms	−0.02 (0.91)	0.11 (0.97)	0.08 (0.99)
Use condoms	0.11 (0.93)	0.28* (1.03)	0.22 (1.13)
Self-talk**	−0.05 (0.62)	0.12 (0.73)	0.21* (0.75)
Psychomotor Skills			
Bleach cleaning	0.89* (1.33)	1.18* (1.29)	1.18* (1.07)
Condom use	0.49* (0.59)	0.63* (0.63)	0.57* (0.51)

[a]The number missing ranges from 0 to 32.
[b]The number missing ranges from 0 to 38.
[c]The number missing ranges from 0 to 36.
*Within-group mean change is statistically significant, $P < .01$.
**Between-group mean changes are statistically significant, $P = .01$.

TABLE 5—Baseline and Follow-up Levels of Composite Drug-Use and Sexual Risk Behaviors in Subjects with Follow-up Information

Risk behavior	Baseline		Follow-up	
	No.	%	No.	%
Drug use (n = 281)				
No injection	64	23	158	56
Injection, no borrowing	56	20	51	18
Borrowing, bleach always used	47	17	34	12
Borrowing, bleach used sometimes	66	23	20	7
Borrowing, bleach never used	48	17	18	6
Sexual[a] (n = 292)				
No vaginal/anal contact	56	19	95	33
1 Partner, condom used always/sometimes	47	16	40	14
1 Partner, condom never used	115	39	101	35
2+ Partners	74	25	56	19

[a]Only penile-vaginal and penile-anal intercourse are considered.

TABLE 6—Odds Ratios (and 99% Confidence Limits) of Reporting Safer Behavior at Follow-up than at Baseline, by Intervention

Behavior	Early Informational	Late Informational	Enhanced
Frequency of injecting	4.7 (2.5,8.6)	3.5 (2.2,5.6)	7.4 (4.3,12.6)
(Given injecting)			
No. sharing partners	4.7 (1.8,12.2)	2.5 (1.0,6.0)	2.1 (1.1,4.1)
Frequency of borrowing	2.3 (1.0,5.1)	2.1 (0.9,4.5)	1.7 (0.8,3.4)
(Given borrowing)			
Frequency of bleaching	1.4 (0.6,3.1)	2.5 (0.6,10.4)	2.1 (0.7,6.3)
Composite drug use	5.2 (2.8,9.4)	2.7 (1.7,4.2)	4.0 (2.5,6.5)
Composite sexual behavior	2.0 (1.1,3.5)	1.5 (0.9,2.4)	1.6 (1.0,2.4)

tions. In addition, other components of the detoxification and rehabilitation program, including pre- and posttest counseling, and the extensive data collection for the research may have contributed to a multiple treatment effect.[12]

There was an ongoing process evaluation component to the evaluation plan. It proved very difficult to monitor how often and with what intensity AIDS issues were discussed in individual or other group therapy sessions, and to assess the impact these encounters may have had. Staff were aware of what intervention was going on at a particular time, and may have consciously or unconsciously provided individual AIDS counseling (compensatory equalization of treatment).[13]

Additional influences outside the program undoubtedly affected behavioral outcomes at follow-up. Knowledge of AIDS and methods of risk reduction have affected risky behaviors. In the Worcester, Mass, area, coordinated efforts to educate intravenous drug users about HIV and to distribute bleach have resulted in significant increases in bleach use.[14] There has also been a significant secular decline in sharing reported by intravenous drug users entering drug abuse treatment.[14] In view of all these changes, the effects of the interventions may have been diluted considerably, and "booster" sessions may be required to maintain any initial effect.

Attrition bias due to differential loss to follow-up in the three intervention groups may have affected the results. However, attrition rates did not differ by intervention group, and except for language, our baseline measurements showed no differences between those followed and not followed. Our attrition rate, 41%, is not surprising, given the difficulties in following drug abusers.[15] Additionally, the wide geographic distribution of residences (less than one third of the clients lived in the city where the treatment program was located) and the fact that the follow-up period of 10 to 18 weeks was deliberately truncated contributed to attrition.

Potential information bias must be acknowledged. The measurement of attitudinal outcomes can be problematic in the evaluation of interventions that aim to change perceptions due to "response shift" bias.[16] Thus, subjects in the enhanced intervention may actually have rated their self-efficacy somewhat lower after the intervention because they had gained a more realistic perspective on the difficulty of carrying out the behaviors.

The finding of decreased perceived susceptibility in all intervention groups was initially viewed as counterintuitive, given the volume of documented studies about the role of education in increasing perceived susceptibility.[17] Participation in the educational intervention may have provided subjects with a more realistic assessment of risk, with a concomitant decrease in perceived susceptibility to AIDS. For example, at baseline there was a tendency for participants to overestimate the actual HIV seroprevalence among intravenous drug users in Massachusetts; estimates were more accurate at posttest.

The educational and policy implications of the study are limited at this early stage of follow-up. However, the finding of no substantial differences in outcome between the early and late informational interventions indicates that there is no benefit to delaying AIDS education, particularly because the longer the delay, the more clients are likely to have left the program prematurely. Evaluations of other "enhanced" models of AIDS education for drug users reported to date have also yielded disappointing results.[18,19] Thus, at this point, research does not support the widespread substitution of enhanced models of AIDS education for drug users in place of simpler, informational models.[20] □

Acknowledgments

This research was supported by grant R01 DA04508 from the National Institute on Drug Abuse.

A preliminary version of this paper was presented at the 118th Annual Meeting of the American Public Health Association, New York, NY, September 30–October 4, 1990.

The authors are indebted to the staff and clients of Spectrum Addiction Services, Inc, for their cooperation with this project.

References

1. Des Jarlais DC, Friedman SR, Casriel C. Target groups for preventing AIDS among intravenous drug users. *J Consult Clin Psychol.* 1990;58:50–56.
2. Becker MH, Rosenstock IM. Social-psychological research on determinants of preventive health behavior. *Behav Sci Prev Med.* 1974;13:173–179.
3. Fishbein M, Ajzen I. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* Reading, Mass: Addison-Wesley; 1975.
4. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice-Hall; 1986.
5. Marlatt GA, Gordon JR. Determinants of relapse: implications for the maintenance of behavior change. In: Davidson P, Davidson S, eds. *Behavioral Medicine: Changing Health Lifestyles.* New York: Brunner/Mazel; 1980:410–452.
6. Marlatt GA, Gordon JR. *Relapse Prevention: Maintenance Strategies in the Treatment of Addiction.* New York: Guilford; 1985.
7. McCusker J, Koblin B, Lewis BF, Sullivan J. Demographic characteristics, risk behaviors, and HIV seroprevalence among intravenous drug users by site of contact: results from a community-wide HIV surveillance project. *Public Health.* 1990; 80:1062–1067.
8. Dixon WJ, Massey FJ Jr. *Introduction to Statistical Analysis.* New York: McGraw-Hill Book Co; 1969.
9. Agresti A. A survey of models for repeated ordered categorical response data. *Stat Med.* 1989;8:1209–1224.
10. SAS institute, Inc. *SAS/STAT User's Guide, Release 6.03 Edition.* Cary, NC: SAS Institute, Inc; 1988.
11. McCusker J, Stoddard AM, Zapka J, et al. Behavioral outcomes of AIDS educational interventions for drug users in treatment. Abstract Th.D. 60, Seventh International Conference on AIDS; June 16–21, 1991; Florence, Italy.
12. Shortell SM, Richardson, WC. *Health Program Evaluation.* St. Louis, Mo: CV Mosby Company; 1978.
13. Cook TD, Campbell DT. *Quasi-Experimentation, Design, and Analysis Issues for Field Settings.* Chicago, Ill: Rand McNally; 1979.
14. McCusker J, Stoddard AM, Koblin BA, Sullivan JL, Lewis BF, Sereti SM. Time trends in risky injection practices in a multisite study in Massachusetts: effects of enrollment site and residence. *AIDS Educ Prev.* In press.
15. Nurco DN, Robins LN, O'Donnell JA. Locating respondents. In: *Conducting Fol-*

low-up Research on Drug Treatment Programs. Rockville, Md: National Institute on Drug Abuse; 1977. Treatment Program Monograph Series No 2:71–84.
16. Howard GS. Response-shift bias: a problem in evaluating interventions with pre/post self-reports. *Eval Rev.* 1980;4(1):93–106.
17. Janz NK, Becker MH. The Health Belief Model: a decade later. *Health Educ Q.* 1984;11:1–47.
18. Chitwood D, McCoy CB, McKay C. Effectiveness of a risk reduction program for intravenous drug users. Presented at the 118th Annual Meeting of the American Public Health Association; September 30–October 4, 1990; New York, NY.
19. Gibson DR, Lovelle-Drache J, Young M, Sorensen JL. Does brief counseling reduce HIV risk in IV drug users? Final results from a randomized clinical trial. Abstract Th.D. 59. Seventh International Conference on AIDS; June 16–21, 1991; Florence, Italy.
20. Dengelegi L, Weber J, Torquata S. Drug users' AIDS-related knowledge, attitudes, and behaviors before and after AIDS educational sessions. *Public Health Rep.* 1990;105:504–510.

APPENDIX—Description of Scaled and Individual Item Variables

Knowledge

Transmission scale

Ranging from 0 to 10, scored from an open-ended question, this scale measures a subject's ability to name the most important ways that HIV is transmitted: via sharing needles without use of bleach, unprotected vaginal and oral sex, and prenatal transmissions.

Prevention of transmission scale

Ranging from 0 to 8, scored from an open-ended question, this scale measures subject's ability to identify methods to prevent HIV, including: not sharing injection equipment, using bleach, using condoms, and limiting sex partners to one.*

Prevention of progression scale

Ranging from 0 to 6, scored from an open-ended question, this scale measures subject's ability to identify ways that a person infected with HIV stay as healthy as possible, including eating well, reducing stress, exercising, getting enough sleep/rest, not using drugs, and avoiding reinfection through not engaging in unsafe sex or unsafe drug use.*

Test knowledge scale

Ranging from 0 to 6, scored from correct response to 5 Yes/No items:
(1) If a person has a positive result on the AIDS antibody test, does it mean that he or she has been infected with the AIDS virus?
(b) Does it mean the person has AIDS?
(c) Does it mean the person will definitely get AIDS sometime in the future?
(d) Does it mean the person can pass the infection on to someone else?
(e) After a person has been infected with the AIDS virus, how soon will he or she usually become positive on the antibody test (right away, within 1 to 3 weeks, within 1–3 months, more than 3 months, never).

Attitudes and Beliefs**

Susceptibility scale

Perceived susceptibility to AIDS measured by 2 items.
Responses: 1 (strongly agree) to 4 (strongly disagree).
Range: 2 to 8. Cronbach's alpha: 0.70.
(a) Your chances of getting AIDS are quite small.
(b) You are quite likely to get AIDS in the next 5 years.

Benefits scale

Perceived benefits of AIDS-related behavior or change measured by 4 items.
Response: 1 (strongly agree) to 4 (strongly disagree).
Range: 4 to 16. Cronbach's alpha: 0.55.
(a) Using a condom during sex would give you peace of mind.
(b) Doing things to protect yourself and others from AIDS would give you a lot more confidence in yourself.
(c) Each time a person refuses to share works, it gets easier the next time.
(d) Each time a person uses a condom during sex, it gets easier to do so again.

Sex slips recovery—single item

Response: 1 (strongly agree) to 4 (strongly disagree).
(a) If a person promises himself or herself to always have "safe sex" and then "slips" once, he or she can still get in control and do better next time.

Bleach slips recovery—single item

Response: 1 (strongly agree) to 4 (strongly disagree).
(a) If a person forgets to use bleach to clean his or her works, he or she can use will power to use bleach the next time.

Self-efficacy (single items)

Confidence to carry out 8 specific behaviors in the next 3 months.
Response choices: 1 (extremely confident), 2 (somewhat confident), 3 (not very confident), 4 (not at all confident).
(a) You could avoid using any drugs.
(b) If you were shooting drugs, you could refuse someone who wanted to share your needles.
(c) If you were shooting drugs and had to share works, you could always clean your works with bleach.
(d) You could always carry bleach with you, "just in case."
(e) You could stay away from people who insist on sharing needles.
(f) You could always carry a condom with you "just in case."
(g) You could always use a condom during sex.
(h) You could "talk yourself out of" doing something which would increase your AIDS risk.

Continued on next page

APPENDIX—Continued.

Psychomotor Skills

Cleaning works with bleach

Range 0 to 3 based on the number of actions completely carried out during demonstration. Each of the following criteria was scored as .5. Completely filling:

(a) the syringe with bleach
(b) the syringe with water the first time
(c) the syringe with bleach the second time
(d) the syringe with water the second time
(e) the cooker with bleach (at least one time)
(f) the cooker with water (at least one time)

Condom use

Range 0 to 1 based on whether respondent pinched the tip of the condom prior to rolling it down on a penis model.

*The detailed algorithm for the weighting of responses can be obtained from the authors
**Scales were computed as the unweighted sum of the items. Item scores were reversed as necessary so that a high score on the scale indicates greater agreement with the construct.

Awards for Excellence in Asthma Control

The First National Conference on Asthma Management is soliciting nominations for awards for excellence in asthma control. The following awards will be presented:

1. *National Asthma Award for Biomedical Research,* awarded for leadership and contributions in research that have led to a better understanding of basic biologic mechanisms in asthma.

2. *National Asthma Award for Clinical Research,* awarded for leadership and contributions in clinical research that have led to major advances in the management of asthma.

3. *National Asthma Award for Health Education Research,* awarded for leadership and contributions in health education research that have emphasized the importance of the patient-physician partnership as the cornerstone of effective asthma management.

4. *National Asthma Award for Public Education/Service,* awarded for leadership in the development of and/or contributions to community asthma programs, patient education, or increased public awareness about asthma.

One award will be presented in each of the above categories. Recipients will be selected by a distinguished awards jury of leaders in the field of asthma research and control. Nominations for the awards must be postmarked by July 1, 1992. For more information, contact the First National Conference on Asthma Management, PO Box 30105, Bethesda, Maryland 20824-0105. Tel: 301/951-3275.

HIV Infection in Intravenous Drug Users Entering Drug Treatment, United States, 1988 to 1989

ABSTRACT

Background. Intravenous drug use has played a key role in the human immunodeficiency virus (HIV) epidemic. Standardized surveillance of HIV infection among intravenous drug users (IVDUs) is needed to determine HIV prevalence rates, to monitor changes in prevalence over time, and to describe behaviors associated with HIV infection.

Methods. In 1987, the Centers for Disease Control began collaborating with state and local health departments to conduct a national program of HIV seroprevalence surveys in a variety of populations and settings. This program includes unlinked HIV seroprevalence surveys in IVDUs entering sentinel drug treatment programs.

Results. From April 1988 through December 1989, annual studies were completed in 59 drug treatment centers in 33 US cities. Center-specific seroprevalence rates ranged from 0% to 48.2%, with a median of 4.6%. HIV seroprevalence rates varied widely by geographic area, with rates highest in the Northeast, intermediate in the Middle Atlantic cities of Baltimore and Washington, DC, and lower in other parts of the country. Median rates were 15.6% among African Americans, 3.2% among Hispanics, and 3.3% among Whites.

Conclusions. Intravenous drug use is likely to remain an important factor in HIV transmission. This study supports the need to develop or expand programs to prevent the further introduction and spread of HIV among IVDUs and to prevent HIV transmission to their sexual partners. (*Am J Public Health.* 1992;82:541–546)

David M. Allen, MD, MPH, Ida M. Onorato, MD, Timothy A. Green, PhD, and the Field Services Branch of the Centers for Disease Control

Introduction

From 1981 through June 30, 1991, the Centers for Disease Control (CDC) received reports of 39 904 cases of acquired immunodeficiency syndrome (AIDS) in heterosexual intravenous drug users (IVDUs), 11 823 cases of AIDS in homosexual and bisexual men who reported intravenous drug use, 5305 cases in heterosexual partners of IVDUs, and 1847 cases in infants born to mothers who were IVDUs or the sexual partners of IVDUs. These cases associated with intravenous drug use represent 32.2% of all AIDS cases.

Because of the lengthy incubation period between the human immunodeficiency virus (HIV) infection and the diagnosis of AIDS,[1] reported AIDS cases do not necessarily represent the current scope or the recent patterns and trends of the HIV infection. In 1987, the CDC reviewed published and unpublished information on HIV seroprevalence among IVDUs in treatment programs in the United States.[2] In the reviewed studies, the HIV seroprevalence rates varied by geographic location, race/ethnicity, and other demographic variables. However, these rates were difficult to interpret because of differences in study methodology.

In 1987, the CDC began collaborating with state and local health departments to conduct a national program of HIV seroprevalence surveys in a variety of populations and settings.[3,4] These surveys, conducted annually, include assessing HIV prevalence among persons attending sentinel drug treatment centers in selected metropolitan areas.

Standardized serological surveillance of IVDUs is needed to determine levels of HIV infection, to monitor changes in prevalence over time, and to describe behaviors associated with HIV infection. Furthermore, these data are needed to project future health care needs and formulate public health policy.

IVDUs entering treatment programs are the most accessible group of all IVDUs for unlinked ("blinded") surveys because, upon entry into drug treatment, many clients receive a routine medical assessment, which entails, among other things, the collection of blood specimens. IVDUs entering drug treatment can serve as an indicator group for monitoring the levels and trends of HIV infection among drug users in the United States. This report describes the serosurveillance system for HIV in IVDUs and presents findings from the first two years of operation.

Methods

Survey methods for this study are described in detail elsewhere.[5] In brief, cities are selected to participate in clinic-based HIV seroprevalence surveys based on (1) the cumulative number of AIDS cases, (2) syphilis and gonorrhea case rates, (3) geographic location, and (4) willingness of

David M. Allen and Ida M. Onorato are with the HIV Seroepidemiology Branch and Timothy A. Green is with the Statistics and Data Management Branch of the Centers for Disease Control (CDC) in Atlanta, GA. Members of the CDC's Field Services Branch are Gary Bledsoe, Dave Collie, Ted Connell, Debbie Deppe, Gary Edgar, Jeff Efird, Willis Forrester, Barbara Meek, Don Ruberti, Ron Sanders, and Steven Sloane.

Requests for reprints should be sent to David M. Allen, MD, MPH, Medical Epidemiologist, Division of HIV/AIDS, Centers for Disease Control, Atlanta, GA 30333.

This paper was submitted to the Journal May 20, 1991, and accepted with revisions September 20, 1991.

state and local health departments to participate. Drug treatment centers within these cities are selected by state and local health departments based on criteria that include the existence of procedures for routine medical assessment of new clients, the characteristics of the populations served, and geographic location. By mid-1990, 114 centers in 44 metropolitan areas had participated in the surveys.

In unlinked surveys, HIV-1 antibody tests are performed on the unused portion of blood specimens collected for other routine purposes. Blood specimens from clients entering treatment programs who have used drugs intravenously in the previous 12 months are eligible for inclusion. General demographic information (sex, age, race or ethnicity, and zip code of residence), as well as the primary intravenous drug used, is recorded for each client on standardized survey forms, each of which is assigned a precoded study number. Self-reported sexual behaviors and sexual orientation are also recorded for clients in centers where this information is routinely collected. Personal identifiers (e.g., client's name, address, and clinic chart number) are removed from the blood specimen, which is labeled with the same study number; only then is the blood specimen tested for antibodies to HIV-1.

Data forms and HIV test results are sent to a designated metropolitan area survey coordinator, who links test results and demographic data. To further ensure anonymity, only summary data—without individual test results—are returned to the clinic.

Blood specimens from eligible clients are sequentially collected until the predetermined survey sample size is achieved. We recommend a sample size of 500 clients per center, or continuous sampling during a 12-month period for centers in which fewer than 500 clients are enrolled. Subsequent annual surveys will be conducted during the same months of the year.

Specimens are tested for antibodies to HIV at designated laboratories using an HIV-1 enzyme immunoassay licensed by the Food and Drug Administration (FDA) and according to manufacturer's recommendations. Specimens that are repeatedly reactive are then tested by an FDA-licensed Western blot assay. Western blot band patterns are interpreted according to recommendations of the Association of State and Territorial Public Health Laboratory Directors and the CDC. Participating laboratories take part in the CDC program of quality assurance.

Participating state and local health departments analyze their data using CDC-designed computer software. Data are also shared with the CDC for regional and national analyses. For each drug treatment center, HIV seroprevalence rates are calculated by demographic (race/ethnicity, sex, age) and other variables, including the primary drug used and treatment modality, for all clients combined. We designated the median among the rates for participating centers in each city as the seroprevalence rate for that city. The median is the point that half of the observations fall above and half fall below.

For this report, we restricted analyses to those centers with 50 or more blood specimens collected from April 1988 through December 1989. Rates by sex, race and ethnicity, age group, and primary drug used were calculated only for those centers with more than 25 specimens for each subgroup. Comparisons of seroprevalence rates were made using the Mantel-Haenzel test.

Results

Through December 31, 1989, 59 drug treatment centers in 33 cities had tested 50 or more eligible samples each. A total of 16 237 blood specimens from these centers were tested for antibodies to HIV-1, of which 2201 were Western blot positive. By center, seroprevalence rates ranged from 0% to 48.2%, with a median rate of 4.6% (Table 1). Six centers reported rates under 1%, and 7 centers reported rates over 30%. Rates varied widely by geographic area, with rates highest in the Northeast, intermediate in the Middle Atlantic cities of Baltimore and Washington, DC, and lower in other parts of the country, including Hawaii (Figure 1). Data were not available from Miami or Puerto Rico. The median HIV seroprevalence rate for centers in the Northeast was 25.8% compared with 2.6% for centers in the remainder of the United States. Within the Northeast, the highest rates were seen in Newark and New York City. Rates decreased to the north and south of these metropolitan areas.

Sex

Of the total study population, 66% were males and 34% were females. Among all centers, the median rate for males was 5.4% (range: 0% to 50.0%) and for females, 4.4% (range: 0% to 50.0%) (Table 1). After controlling for differences among centers and age groups, there were no statistically significant differences between male and female seropositivity rates. In selected cities, especially those with high seroprevalence rates (i.e., Boston, Camden, Baltimore, and Washington, DC), rates were higher in women than in men; these differences, however, were not statistically significant after controlling for age.

Race

Table 2 lists HIV seroprevalence rates by race/ethnicity by city. For all centers, median rates were 15.6% for African Americans, 3.2% for Hispanics, and 3.3% for Whites. African Americans had a significantly ($P < .001$) higher seroprevalence rate than Whites after controlling for center. Among those 33 centers in 22 cities with both African-American and White drug users tested, the median rates were 16.9% and 4.3%, respectively. Hispanics had a significantly ($P < .002$) higher seroprevalence rate than Whites after controlling for center. For those 19 centers in 12 cities with both Hispanic and White drug users tested, the median rates were 3.4% and 4.2%, respectively. Although the median rates are similar, the rate among Hispanics exceeds 15% in eight centers and 30% in five centers, whereas the rate among Whites exceeds 15% in six centers and 30% in only one center.

Age

Preliminary graphic analysis indicated that the effect of age on HIV seroprevalence rates is different for high versus low seroprevalence centers. As a result, we categorized centers into three groups based on HIV seroprevalence rates. Group 1 (n = 28) included centers with HIV seroprevalence rates from 0.0% to 9.9%, group 2 (n = 12) included centers with rates from 10.0% to 21.5%, and group 3 (n = 7) included centers with rates of 21.6% and above. Figure 2 shows the relationship between seroprevalence rate and age group. After controlling for center there were significant age-group effects, with HIV seroprevalence rates increasing with age through age 40 in groups 2 ($P < .01$) and 3 ($P < .001$) but not in group 1. When the 20- to 24-year-old age group was omitted from the analysis, however, the age-group effect in group 2 was only marginally significant ($P < .04$).

Primary Drug and Treatment Modality

Heroin was the primary drug used for 13 499 clients in 52 centers, and cocaine was the primary drug for 899 clients in 9

centers. Median center-specific seroprevalence rates for heroin users was 5.4% (range: 0.0% to 48.2%) and for cocaine users, 8.3% (range: 1.1% to 18.8%).

Twenty-two cities had centers where clients admitted to treatment listed heroin alone as the primary drug used. A small number of clients chose other drugs as the primary drug used. Five cities had centers that permitted the comparison of heroin with cocaine as the primary drug used. In none of these centers were the differences in seroprevalence rates statistically significant.

Methadone detoxification or maintenance was the treatment modality for most clients (Table 3). HIV seroprevalence was higher in clients entering cocaine and drug-free treatment compared with those entering methadone detoxification or methadone maintenance. These rates, however, are based on small sample sizes.

Sexual Orientation

Of all the centers, three collected information on the sexual orientation of new clients and included 10 or more homosexual or bisexual men or women. The three centers include a total of 61 bisexual or homosexual women, 55 bisexual or homosexual males, 346 heterosexual women, and 764 heterosexual males. HIV seroprevalence rates for homosexual or bisexual men in these three centers were 21.4%, 21.7%, and 44.4%, compared with 4.3%, 14.0%, and 3.4% for heterosexual men in the same centers. These differences were statistically significant in two of the three centers ($P < .05$). Rates for homosexual or bisexual women were 0.0%, 0.0%, and 3.0%, compared with 2.9%, 1.1%, and 2.1% for heterosexual women in the same clinics. These differences were not statistically significant.

TABLE 1—HIV Seroprevalence Rates in Drug Treatment Center Entrants by Metropolitan Area and Sex, April 1988 through December 1989

Region/Metropolitan Area	No. of Centers	No. of Sera Tested	Median[a] % positive		All centers % positive
			Male	Female	Median (Range)[a]
Northeast					
Boston	3	944	13.6	19.4	15.9 (5.2–32.1)
Providence	1	131	21.6	14.3	19.1
New York City	3	1724	39.9	24.8	33.6 (19.6–44.7)
Newark	2	404	44.3	44.6	43.4 (38.6–48.2)
Atlantic City	1	403	17.8	16.4	17.4
Camden	1	89	8.2	17.9	11.2
Paterson	1	345	40.8	30.4	36.8
New Brunswick	1	143	20.8	14.9	18.9
Jersey City	1	125	41.5	24.1	32.8
West					
Denver	2	313	1.2	2.6	1.8 (1.4–2.2)
Phoenix	2	497	2.9	5.0	3.5 (2.6–4.4)
Salt Lake City	1	207	3.0	1.4	2.4
Seattle	5	1390	1.3	0.0	0.9 (0.0–4.3)
Portland	1	56	0.0	—	0.0
San Francisco	3	589	6.3	7.6	7.9 (2.9–15.0)
Los Angeles	5	1621	3.5	2.3	3.0 (1.0–4.6)
Honolulu	1	288	2.1	1.1	1.7
Albuquerque	1	62	0.0	—	0.0
North Central					
Chicago	1	1179	13.5	12.8	13.2
Cleveland	2	424	5.1	4.0	5.0 (4.7–5.2)
Detroit	1	369	7.8	5.1	7.0
St. Louis	1	50	0.0	—	0.0
South					
Baltimore	4	1151	16.2	19.3	17.8 (15.3–18.6)
Washington, DC	1	1393	17.0	21.8	18.6
Richmond	1	91	3.2	0.0	2.2
Atlanta	2	491	10.7	11.7	10.8 (6.4–15.2)
Jacksonville	1	103	9.6	3.3	7.8
Memphis	1	125	3.8	4.3	4.0
Birmingham	2	147	4.6	—	4.0 (1.1–6.9)
New Orleans	2	255	1.0	2.7	1.6 (1.6–1.6)
Dallas	1	446	2.2	2.4	2.2
Houston	3	383	1.9	2.9	2.0 (1.0–2.2)
Oklahoma City	1	299	2.1	0.0	1.7
All areas studied	59	16 237	5.4	4.4	4.6 (0.0–48.2)

[a]The median rate of all centers in the metropolitan area; range reflects the lowest and highest rates of all centers in the metropolitan area.

Discussion

Data from the surveys of IVDUs entering treatment indicate that HIV seroprevalence is strongly associated with geographic location, with rates highest in the Northeast. This geographic pattern may be due to several reasons, including the earlier entry of HIV in intravenous drug-using populations in the Northeast; differences in both the availability of sterile injection equipment and the prevalence of drug use behaviors that facilitate HIV transmission, such as the use of "shooting galleries";[6] and differences in the development and characteristics of HIV prevention programs. The reasons for higher HIV prevalence in IVDUs in the Northeast and lower prevalence in the remainder of the United States, however, require further clarification.

HIV seroprevalence rates in IVDUs also vary by race and ethnicity, with rates higher in African Americans and Hispanics than in Whites. This pattern is consistent with that seen for drug use–associated AIDS.[7] The magnitude of the differences in HIV seroprevalence between African Americans and Whites was inconsistent, however, and varied by city.

There were no significant differences in HIV seroprevalence rates in IVDUs by sex. Although this finding is in contrast to differences in the proportion of AIDS cases by sex, HIV seroprevalence studies of IVDUs, especially those in high-prevalence areas, have shown similar prevalence rates in males and females.[8,9]

In centers with HIV seroprevalence of at least 10%, rates increased with age. This may be due to several reasons, including the risk of HIV in clients entering treatment with a longer history of drug use compared with younger clients, or the aging of a cohort of IVDUs that was previously infected with HIV.

Self-reported information on sexual behaviors and sexual orientation is not routinely collected upon entry into drug treatment. In the three centers where this information is collected, the data indicate that bisexual or homosexual male IVDUs have higher HIV seroprevalence than heterosexual male IVDUs. Although the data are limited, they suggest that sexual ori-

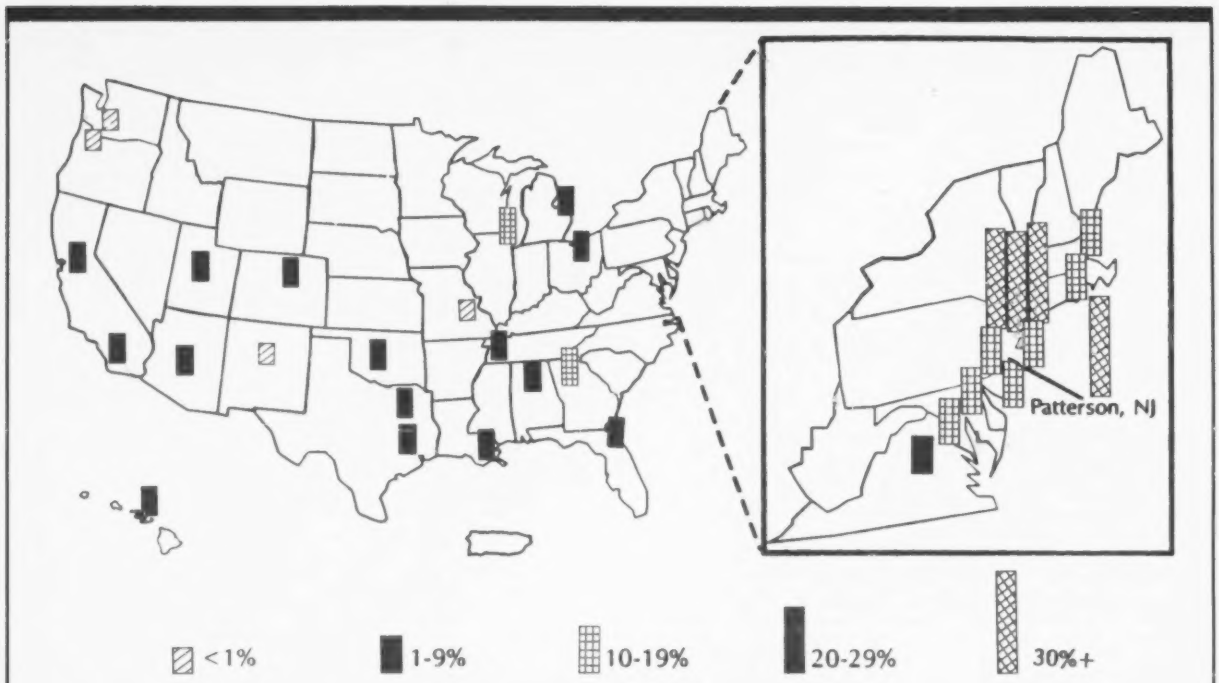


FIGURE 1—Median HIV seroprevalence in drug treatment centers, United States, 1988–1989.

entation and sexual behaviors should be ascertained as part of the routine clinical assessment of the risk of HIV for persons entering drug treatment. Men who report sexual activity with men may require additional or specialized HIV prevention efforts.

There are several considerations in the interpretation of these data. First, the seroprevalence rates, although indicative of patterns throughout the country, are truly representative of only those centers studied. This survey, however, contains the only national collection of blood specimens with standardized protocols for IVDUs entering treatment. Second, because participating clinics do not constitute a probability sample of all drug treatment facilities (nor is such a sample feasible), the data are illustrative rather than representative of HIV seroprevalence in IVDUs entering drug treatment throughout the United States or in a given metropolitan area. Third, the study includes sera only from IVDUs entering drug treatment and not from those who do not seek treatment. While these biases make it difficult to generalize sentinel clinic data to all IVDUs in an entire metropolitan area, repeating the surveys annually should keep most clinic-specific biases constant over time. The resulting data should reflect HIV seroprevalence trends over time for these clinics and assist locally in the development and evaluation of prevention programs.

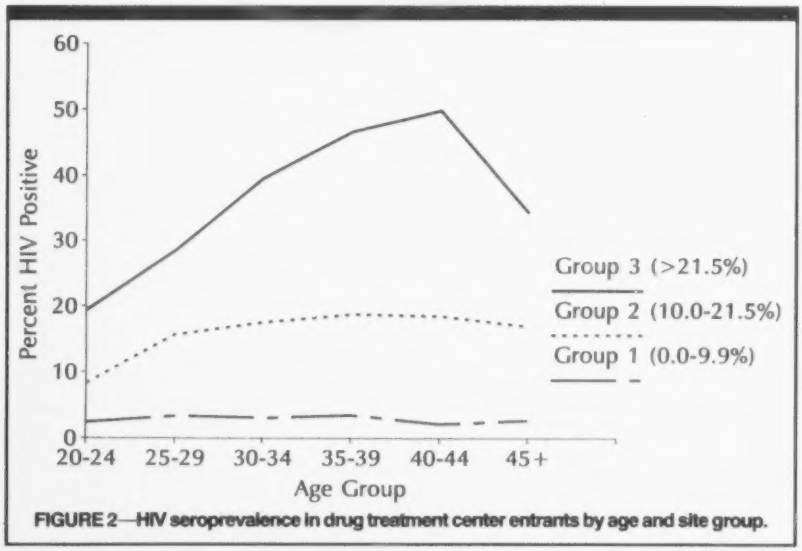


FIGURE 2—HIV seroprevalence in drug treatment center entrants by age and site group.

Current federal requirements stipulate that a routine medical assessment that includes the collection of blood specimens take place upon a person's entry to methadone treatment. Other treatment programs, including those for cocaine or amphetamine use, may not require blood collection prior to treatment. Drug treatment centers that do not routinely collect blood during client admission are not eligible to conduct unlinked surveys. As a result of these requirements, most of the participating centers represented in this report offer methadone treatment. HIV seroprevalence rates in clients receiving

other treatment modalities are therefore less adequately accounted for here.

Limited demographic and risk behavior data are available in unlinked surveys. These surveys indicate the need for additional information on factors associated with the initiation of drug use; behaviors associated with HIV infection, such as the frequency of drug use; and the possible reasons for increased rates in certain geographic areas and among selected racial and ethnic minorities.[10] In particular, higher rates among selected racial and ethnic minority groups compared with rates among Whites may be mediated by other factors, such as socioeconomic status or geographic location within metropolitan areas. Selected drug treatment centers conducting unlinked seroprevalence surveys also conduct risk behavior surveys in which consenting clients are interviewed in detail regarding potential risk behaviors. These linked surveys will complement the unlinked studies and may suggest further preventive measures aimed at changing high-risk behaviors and reaching the sexual partners of IVDUs who may be at risk of HIV infection.

Intravenous drug use has been associated with explosive increases in HIV transmission in the New York City area in the early 1980s[11] and more recently in parts of Europe[12] and Asia.[13] Because the potential for such an explosion also exists in parts of the United States where HIV prevalence is low, sentinel surveys provide an important early warning system to indicate changes in HIV prevalence among drug users throughout the United States. In clinics offering voluntary HIV testing and counseling, unlinked surveys can also be used to indicate the degree to which HIV-infected clients are included in testing and counseling programs.

HIV transmission in IVDUs, their sexual partners, and their babies remains an important public health problem, and continued efforts should be made to develop programs designed to prevent the transmission of HIV to uninfected IVDUs and to the sexual partners of IVDUs. Intravenous drug use is likely to remain the most critical factor associated with HIV in women and in heterosexual and perinatal transmission. Use of sentinel HIV surveillance data can assist in developing programs to prevent the introduction and spread of HIV among IVDUs and their sexual partners. □

TABLE 2—HIV Seroprevalence Rates in Drug Treatment Center Entrants by Metropolitan Area and Race/Ethnicity, April 1988 through December 1989

Region/Metropolitan Area	Median % positive[a]					
	No. of Centers	African American	No. of Centers	White	No. of Centers	Hispanic
Northeast						
Boston	2	39.6	3	10.4	1	43.7
Providence	—	—	1	18.6	—	—
New York City	3	41.4	3	26.4	3	33.3
Newark	2	46.6	1	25.6	1	41.3
Atlantic City	1	19.9	1	13.5	1	17.4
Camden	—	—	1	0.0	—	—
Paterson	1	57.4	1	23.7	1	30.4
New Brunswick	1	18.2	1	16.2	—	—
Jersey City	1	18.8	1	41.8	—	—
West						
Denver	—	—	2	3.3	2	0.4
Phoenix	—	—	2	4.3	2	2.5
Salt Lake City	—	—	1	1.7	—	—
Seattle	4	1.6	5	1.2	—	—
San Francisco	1	28.2	3	2.1	1	3.4
Los Angeles	2	8.1	4	2.9	4	0.8
Honolulu	—	—	1	1.9	—	—
Albuquerque	—	—	—	—	1	0.0
North Central						
Chicago	1	13.5	1	11.4	1	17.5
Cleveland	2	5.4	2	2.9	—	—
Detroit	1	8.8	1	1.3	—	—
St. Louis	1	0.0	—	—	—	—
South						
Baltimore	4	21.3	4	7.1	—	—
Washington, DC	1	19.5	1	9.4	—	—
Richmond	1	3.2	1	0.0	—	—
Atlanta	2	13.9	2	5.8	—	—
Jacksonville	—	—	1	5.3	—	—
Memphis	1	6.3	1	2.7	—	—
Birmingham	1	14.3	2	0.8	—	—
New Orleans	1	1.2	2	1.3	—	—
Dallas	1	4.3	1	0.5	1	0.0
Houston	—	—	3	1.2	2	1.6
Oklahoma City	1	2.2	1	1.7	—	—
All areas studied	36	15.6	54	3.3	21	3.2

[a]The median rate of all the centers in the metropolitan area.

TABLE 3—HIV Seroprevalence Rates in Drug Treatment Center Entrants by Treatment Modality, April 1988 through December 1989

Treatment Modality	No. of Centers	No. of Sera Tested	Median[a] % positive	(Range)
Methadone maintenance	50	8942	5.7	(0.0–55.1)
Methadone detoxification	31	4299	5.8	(0.0–38.9)
Cocaine	2	513	12.0	(8.9–15.2)
Drug free	4	295	12.6	(0.0–21.3)
Therapeutic community	5	536	9.8	(0.0–15.9)
Other	5	368	6.1	(5.3–42.4)

[a]The median rate of all the centers within each treatment modality; range is the lowest and highest rates.

Acknowledgments

We are grateful to state and local health departments who participated in this national program. We also acknowledge the contributions of Carolyn Sanchez for software development; Van Munn for data management; and other members of the Clinic and Special Surveys Section, Division of HIV/AIDS, for analytical and program support: Eugene McCray, MD, Patricia Sweeney, MPH, Robert Hnath, MPH, Susan Danner, Deborah Wendell, MPH, and Cathy Spruill. We also thank Timothy J. Dondero, MD, Lawrence L. Sanders, MD, MBA, and T. Steven Jones, MD, for comments on this manuscript.

References

1. Medley GF, Anderson RM, Coc DR, Billard L. Estimating the incubation period for AIDS patients. *Nature.* 1987;328:719–721.
2. Hahn RA, Onorato IM, Jones TS, Dougherty J. Prevalence of HIV infection among intravenous drug users in the United States. *JAMA.* 1989;261:2677–2684.
3. Dondero TJ, Pappaioanou M, Curran JW. Monitoring the levels and trends of HIV infection: the Public Health Service's HIV Surveillance Program. *Public Health Rep.* 1988;103:213–220.
4. Pappaioanou M, Dondero TJ, Peterson LR, Onorato IM, Sanchez CD, Curran JW. The family of HIV seroprevalence surveys: objectives, methods, and uses of sentinel surveillance for HIV in the United States. *Public Health Rep.* 1990;105(2):113–119.
5. Jones TS, Allen DM, Onorato IM, Petersen LR, Dondero TJ, Pappaioanou M. HIV seroprevalence surveys in drug treatment centers. *Public Health Rep.* 1990;105(2):125–130.
6. Schoenbaum EE, Hartel D, Selwyn P, et al. Risk factors for human immunodeficiency virus infection in intravenous drug users. *N Engl J Med.* 1989;321:874–879.
7. Centers for Disease Control. Update: acquired immunodeficiency syndrome (AIDS) associated with intravenous drug use, United States, 1988. *MMWR.* 1989;38:165–170.
8. Des Jarlais DC, Friedman SR. Gender differences in response to HIV infection. In: Bridge PT, ed. *Psychological, Neuropsychiatric, and Substance Abuse Aspects of AIDS.* New York, NY: Raven Press; 1988:159–163.
9. Battjes RJ, Pickens RW, Amsel Z, Brown LS. Heterosexual transmission of human immunodeficiency virus among intravenous drug users. *J Infect Dis.* 1990; 162:1007–1011.
10. Krueger LE, Wood R, Dier P, Maxwell C. Poverty and HIV seropositivity: the poor are more likely to be infected. *AIDS.* 1990;4(8):811.
11. Des Jarlais DC, Friedman SR, Novick DM, et al. HIV-1 infection among intravenous drug users in Manhattan, New York City, from 1977 through 1987. *JAMA.* 1989;261:1008–1012.
12. Des Jarlais DC, Friedman SR. HIV infection among intravenous drug users: epidemiology and risk reduction. *AIDS.* 1987;1:67–76. Review.
13. Uneklabh C, Phutiprawan T, Uneklabh T. Prevalence of HIV infection among Thai drug dependents. Presented at the Fourth International Conference on AIDS; June 15, 1988; Stockholm, Sweden.

International Lactation Consultant Association Announces 1993 Manuscript Competition

The International Lactation Consultant Association is sponsoring a 1993 competition designed to identify outstanding manuscripts relating to human lactation or breastfeeding. Authors may enter one of two categories: a literature review competition or a research paper competition. The winner in each category will receive $500 (US). Winning manuscripts will be published in the *Journal of Human Lactation.* Entrants need not be members of the International Lactation Consultant Association. Students are encouraged to enter.

Competition in the research paper competition is open to published or unpublished authors; authors entering the literature review competition is limited to currently unpublished authors. The postmark deadline for entry in the 1993 competition is January 5, 1993.

To obtain specific information regarding this manuscript competition, send a self-addressed (and stamped if US) envelope to K. G. Auerbach, PhD, IBCLC, Editor-in-Chief, Journal of Human Lactation, 1993 Manuscript Competition—Dept. AP, 2240 Willow Road, Homewood, IL 60430-3221, USA.

HIV Prevalence and Risk Behavior among Injecting Drug Users Who Participate in "Low-Threshold" Methadone Programs in Amsterdam

ABSTRACT

Introduction. After the onset of the HIV epidemic, the hope arose that the "harm reduction" approach of the Amsterdam "low-threshold" methadone programs would have a function in controlling the spread of HIV among injecting drug users (IDUs).

Methods. To find evidence for this assumption, 386 IDUs participating in an HIV study between 1985 and 1989 were studied. Long-term regular participants in low-threshold methadone programs were compared with short-term and/or irregular participants.

Results. Controlling for possible confounders, long-term regular participants had a slightly increased but not statistically significant risk of being HIV infected. With regard to current drug use, they injected daily and shared needles as often as did short-term and/or irregular participants.

Conclusions. The results suggest that the Amsterdam low-threshold methadone programs fail to diminish HIV risk behavior. If these programs want to have a positive effect on the spread of HIV, expansion of existing HIV prevention efforts, an increase in the daily methadone dose level, and measures to enforce regular attendance may be required. (*Am J Public Health.* 1992;82:547–551)

Christina Hartgers, MS, Anneke (J.A.R.) van den Hoek, MD, PhD, Pieta Krijnen, MS, and Roel A. Coutinho, MD, PhD

Introduction

Methadone treatment for intravenous heroin users can help reduce the risk of infection with the human immunodeficiency virus (HIV). The magnitude of this reduction depends on the ability of methadone programs to reduce illicit opiate injecting.[1–5]

The Amsterdam "low-threshold" methadone programs endeavor to take down the barriers that treatment programs and/or methadone maintenance programs[6,7] put up for drug users (DUs) by not prohibiting illicit use of opiates or other drugs. This enables these programs to contact many active DUs who do not want to quit drug use. Low-threshold programs do not aim at treating the addiction; instead, they aim primarily at contacting heroin users and regulating or stabilizing their use by preventing withdrawal symptoms. Within the "harm reduction" approach,[8,9] the programs' function is to signal and take care of sociomedical problems in an early phase. Together with social support, this approach should provide a foundation on which initiatives to quit drug use will have a better chance of succeeding.[10]

After the onset of the HIV epidemic, the hope arose that low-threshold methadone programs would have a function in controlling the spread of HIV,[9,11] either by informing active DUs about the HIV risk or through the prescription of methadone, which would stabilize addiction and lifestyle and thereby reduce injecting. The present study aims at evaluating the effect of the Amsterdam low-threshold programs on HIV risk behavior and on the spread of HIV among injecting drug users (IDUs) participating in an epidemiological study of HIV infection.

Methods

Low-Threshold Methadone Programs

An estimated 5000 to 6000 DUs were in Amsterdam in each quarter of 1985 to 1989.[12,13] Among them, approximately 40% had injected in the previous month. The Drugs Department of the Municipal Health Service operates seven low-threshold methadone clinics,[12–14] which daily provide on-site primary medical and social care and supply methadone to a yearly population of about 3500 DUs. Nonresidents get methadone only when it is medically indicated or when it is needed to enable the user to return to the city or country of origin.

Illicit opiate or cocaine use is not a reason for dismissal from the program. Neither is discontinuation of the program for several days, although a reintake (which entails seeing a doctor) is then required. If clients want to quit drug use, they either enter a methadone reduction scheme with urine screening for opiates or are directed to a treatment clinic. Regulated DUs get methadone prescribed by their general practitioner. Between 1981 and 1984, the flow to general practitioners was approximately as large as the opposite flow.[14]

The authors are with the Municipal Health Service, Department of Public Health and Environment, Amsterdam, the Netherlands.

Requests for reprints should be sent to Christina Hartgers, Municipal Health Service, Department of Public Health and Environment, PO Box 20244, 1000 HE Amsterdam, the Netherlands.

This paper was submitted to the Journal July 8, 1991, and accepted with revisions November 27, 1991.

TABLE 1—Logistic Regression Model of Predictors of HIV Infection among Amsterdam IDUs[a]

	Adjusted Odds Ratio (n = 377)	95% Confidence Interval
First occurrence of injecting		
<5 years ago	1	
≥5 years ago	2.39	1.27–4.47
Last occurrence of injecting		
≥6 months ago	1	
<6 months ago	1.91	0.98–3.75
Frequency of borrowing[b]	1.21	0.99–1.48
Mainly heroin and cocaine injecting in the previous 5 years		
No	1	
Yes	2.13	1.30–3.49
German or South European[c] nationality	3.53	1.61–7.77
Other, incl. Dutch	1	
Time living in Amsterdam[d]	1.42	1.20–1.68

[a]Only IDUs that had no missing values on any of the variables used in the analysis (n = 377) were included.
[b]Values: 0 = never, 1 = 1 time, 2 = 2–9 times, 3 = 10–99 times, 4 = 100 times or more; a value of x is associated with a risk of being HIV seropositive that is 1.21 times higher than a value of x–1.
[c]French, Spanish, Italian, Greek, Portuguese, or Yugoslav.
[d]Values: 1 = no time or <1 year, 2 = 1–<2 years, etc., to 6 = ≥5 years; a value of x is associated with a risk of being HIV seropositive that is 1.42 times higher than a value of x–1.

In 1987, the average daily methadone dose in the low-threshold programs was approximately 35 mg (range: 5 to 120 mg). Doses between 60 and 120 mg were prescribed in about 3% of cases. The average yearly duration of methadone supply for Dutch DUs was 21 weeks.

Study Population and Variables

In 1985 an epidemiological study of HIV infection among DUs[15–17] was initiated. This ongoing study involves voluntary, confidential HIV antibody testing and counseling for DUs, combined with an interview by specially trained nurses using a standard demographic/behavioral questionnaire. DUs enroll in the study through a low-threshold methadone clinic or through a sexually transmitted diseases clinic for addicted prostitutes. Most DUs entering the study (83%) have at some time injected; 55% have injected in the month before entering.

The present study concerns all 386 DUs who had injected drugs at some time (hereafter called "injecting drug users," or IDUs) and who enrolled in the study through low-threshold methadone clinics from December 1985 until March 1989.

Interviews and serological tests were completed for all 386 IDUs; 110 (28%) were HIV seropositive. The mean age of the 386 IDUs at entrance was 29.7 years (standard deviation = 5.9, range: 16 to 57); 260 were male (67%) and 126 female (33%); and 311 (81%) were Dutch.

Limited data are available on DUs who get methadone on prescription in Amsterdam[12,13] but are not available specifically on IDUs. With regard to age, there was no significant difference between the 386 study participants and all DUs who got methadone on prescription in Amsterdam in 1987. The male/female ratio in the study was 2:1, as compared with 3:1 among DUs attending low-threshold methadone clinics in 1987.

Behavior reported over the 6 months preceding the interview is called "current" behavior. Behavior in the period prior to these 6 months as far back as 5 years before the interview is indicated as occurring "in the previous 5 years." Current IDUs who started injecting at least 5 years ago are called "long-term IDUs." The variable frequency of borrowing used needles concerns the full 5 years preceding the interview. Long-term regular methadone users (LTM users) were defined as IDUs who started using methadone at least 5 years preceding the interview and who reported daily methadone use in the 5 years preceding the interview. The non–LTM users were defined as IDUs who started using methadone less than 5 years preceding the interview and IDUs who reported irregular use of methadone during these 5 years. Time of the interview was divided into four consecutive periods of 10 months each: (1) December 1985 to September 1986, (2) October 1986 to July 1987, (3) August 1987 to May 1988, and (4) June 1988 to March 1989.

Statistical Analysis and Serology

Statistics used were χ^2, Student's t, Spearman rank correlation coefficient, Mann-Whitney test, odds ratios (ORs), and 95% confidence intervals (CIs).[18] P values (two sided) of less than .05 were taken as significant. Logistic regression models were used to determine the independent contribution of different characteristics in predicting HIV serostatus. All demographic and drug use variables that had a univariate statistically significant association with HIV serostatus were considered for entry into the model. Three injecting variables (first and last occurrence of injecting and frequency of borrowing) were considered crucial and were entered at the first step. At each further step, the variable that contributed most to the prediction of HIV serostatus was selected. To provide estimates of the statistical association between LTM use and HIV infection while controlling for possible confounders, adjusted ORs and CIs were derived from logistic regression coefficients. Serological tests used for HIV testing were enzyme-linked immunosorbent assays and/or immunoblotting.[16]

Results

Predictors of HIV Infection

A model was built (see Methods) that optimally could predict HIV serostatus. First occurrence of injecting, injecting mainly heroin and cocaine in the previous 5 years, nationality, and time living in Amsterdam were significantly associated with HIV infection (see Table 1).

Differences between LTM Users and Non–LTM Users

One hundred and ninety four IDUs (50%) met the criteria for LTM use and 189 (49%) fell in the non–LTM group (3 IDUs had missing data). Table 2 shows the differences between the two groups with regard to duration and regularity of methadone use. Compared with non–LTM users on demographic and drug use variables, LTM users are older, more often male, and more often Dutch; they have lived longer in Amsterdam, started injecting longer ago, and inject mainly heroin less often and cocaine, either by itself or together with heroin ("speedball"), more often (see Table 3).

LTM users inject daily as often as non–LTM users. Among the 189 non–LTM users, 135 persons (71%) currently use methadone daily and 54 persons (29%) do not (see Table 2). The proportion of current daily injectors is also similar in these subgroups: 49 of 135 (36%) versus 19 of 54 (35%). Frequency of borrowing and reuse of own needles and syringes are similar among LTM and non–LTM users (see Table 3).

LTM users have a higher HIV prevalence than non–LTM users: 37% versus 20% (OR = 2.42, CI = 1.54–3.83). There is a positive relation between number of years since the first methadone prescription and being HIV positive ($P = .005$). The higher seroprevalence of LTM users is found in each of the four consecutive time periods: for LTM users it is 35%, 49%, 28%, and 32%, respectively, as compared with 21%, 16%, 24%, and 18% among non–LTM users.

Controlling for Possible Confounding

To control for possible confounding, the variable of LTM use was entered in a logistic regression model together with the variables of age, sex, German or South European nationality, time living in Amsterdam, first occurrence of injection, and last occurrence of injection, with HIV serostatus as a dependent variable. The adjusted OR was 1.60 (CI = 0.93–2.74), indicating a slightly—but not statistically significant—increased risk of being HIV seropositive for LTM users.

As we considered duration of injecting an important possible confounder, we explored the relation between LTM use, HIV risk behavior, and HIV serostatus separately in the subgroup of 223 long-term IDUs. Among these IDUs, duration and regularity of methadone use in the 133 LTM users and 90 non–LTM users were similar to what was found in the total sample (see Table 2). With regard to demographic and drug use variables, the differences (and similarities) between the 133 LTM users and 90 non–LTM users were the same as in the total group, with one exception: the proportion of males among LTM users and non–LTM users are the same. As in the total sample, the 133 LTM users are more often HIV positive than the 90 non–LTM users (43% vs 28%; OR = 1.97, CI = 1.10–3.52). Adjusted for age, sex, nationality, and time living in Amsterdam, the OR for LTM use was 1.34 (CI = 0.69–2.63).

To exclude possible confounding due to exposure to HIV before coming to Amsterdam, we did a separate analysis among Dutch IDUs who were Amsterdam residents for at least 5 years. In this subgroup of 230 IDUs, the 147 LTM users differ from the 83 non–LTM users on four demographic and drug use variables: LTM users are older, started injecting longer ago, currently inject more often, and inject mainly cocaine more often. With regard to current daily injecting, having borrowed at least 10 times, and current borrowing, there are no significant differences between LTM users and non–LTM users. LTM users are again more often HIV positive than non–LTM users (38% vs 21%; OR = 2.39, CI = 1.22–4.71). Adjusted for age, sex, and first and last occurrence of injecting, the OR was 2.06 (CI = 1.02–4.16).

HIV infection may have been the reason for long-term participation in low-threshold methadone programs. To exclude this potential bias, we added the variable daily methadone use to a separate multivariate logistic regression analysis of incident HIV infections.[19] The behavior of 31 IDUs, who were HIV-seronegative at entrance into the study but who seroconverted for HIV during follow-up, was compared with that of 202 randomly selected controls, who remained seronegative throughout their follow-up period. In

TABLE 2—Duration and Regularity of Methadone Use of LTM users and Non–LTM users among Amsterdam IDUs

	LTM users (n = 194)	Non–LTM users (n = 189)
Mean number of years (SD) since first methadone prescription	8.4 (3.2)	2.9 (2.6)
(Median, range)	(7, 5–20)	(2, 0–17)
Daily methadone use in the previous 6 months (%)	100	71
Daily methadone use in period before this (for LTM users, up to 5 years before the interview; for non–LTM users, since first methadone prescription, no more than 5 years preceding intake) (%)	100	60

TABLE 3—Demographic and Drug Use Characteristics of LTM users versus Non–LTM users among Amsterdam IDUs (n = 383[a])

Characteristics	LTM users (n = 194)	Non–LTM users (n = 189)	Significance[b]
Mean (median) age (y)	31.0 (31)	28.3 (28)	$P < .001$
	# %	# %	
Male sex	145 (75)	113 (60)	$P < .002$
Dutch nationality	172 (89)	136 (72)	$P < .0001$
Living >2 years in Amsterdam	181 (93)	127 (67)	$P < .0001$
First time injected ≥5 years ago	162 (84)	118 (62)	$P < .0001$
Currently injecting (<6 months ago)	158 (81)	139 (74)	ns
Currently injecting			
mainly heroin	22 (11)	39 (21)	$P < .01$
mainly cocaine	17 (9)	14 (7)	ns
mainly heroin and cocaine	85 (44)	67 (35)	ns
Injecting in previous 5 years			
mainly heroin	44 (23)	66 (35)	$P < .008$
mainly cocaine	14 (7)	5 (3)	$P < .04$
mainly heroin and cocaine	88 (45)	63 (33)	$P < .02$
Currently daily injecting	75 (39)	68 (36)	ns
Daily injecting in previous 5 years	139 (72)	130 (69)	ns
Currently having borrowed	80 (42)	73 (39)	ns
Having borrowed ≥10 times in previous 5 years	73 (38)	67 (36)	ns
Currently reusing same needle and syringe	82 (42)	81 (43)	ns

[a]As a result of missing data, the effective total sample varies from 378 to 383 subjects. Percentages are calculated with the effective sample as denominator.
[b]Statistics used: to test for differences in mean age, *t* test; for other variables, χ^2 (with $df = 1$) or Mann-Whitney test. Ns = not statistically significant.

univariate analysis, daily methadone use, reported over a mean period of 8 months preceding the first visit after seroconversion or the selected control visit, was not "protective" (OR = 0.67, CI = 0.31–1.47). When the variable daily methadone use was adjusted for the independent risk factors for HIV seroconversion, its OR became 0.81 (CI = 0.35–1.83).

Discussion

An important finding of the present study is that IDUs participating in low-threshold programs who report long-term daily use of methadone engage in similar or even more HIV risk behavior than IDUs who use methadone for a shorter time and/or irregularly. Univariately, being HIV seropositive is associated with LTM use and with the number of years since the first methadone prescription. In contrast, studies from other countries have reported that DUs in long-term methadone maintenance have a lower HIV seroprevalence than DUs not in treatment,[20-24] which suggests that these maintenance programs are successful in protecting IDUs from HIV infection. In the present study, the HIV prevalence of LTM users who entered the study before October 1986 (i.e., in the first time period) is 35%. These IDUs started using daily methadone before 1981. Because HIV prevalence among 145 DUs entering low-threshold programs in 1983 and 1984 was only 3.4%,[25] presumably HIV hardly circulated among DUs in 1981. This indicates that these LTM users got infected mainly while enrolled in the low-threshold programs.

The present study has some limitations. First, the short average duration of yearly methadone supply in low-threshold methadone programs (21 weeks) suggests that, in reality, LTM users did not use methadone continuously in the 5 years preceding intake. Due to privacy regulations, it is impossible to validate our data with data from the methadone program records. Thus, it is unknown whether the actual yearly duration of methadone use in the total sample significantly differs from our data. Nevertheless, we hold it possible that the self-reports overestimate the actual duration of methadone use. For the present study, it is assumed that LTM users were supplied with methadone more continuously—and during a much longer time—than non–LTM users.

Second, because no data on IDUs in low-threshold programs are available, the representativeness of the sample is unknown. However, although our sample is not atypical for low-threshold program participants with regard to number of years of methadone use[13,26] or age, males are underrepresented in the present study.

Third, comparing methadone users with non-users, preferably randomly assigned, would have been more appropriate. However, to find a group of comparable IDUs who have never used methadone is—within the Amsterdam context—impossible: only young or newly starting injectors have never used methadone, and so random assignment was not feasible.

Finally, HIV-seropositive IDUs may have entered the study preferentially—for example, because of health problems. However, all IDUs in the present study entered before March 1989, when zidovudine or prophylaxis for Pneumocystis carinii pneumonia or tuberculosis was not yet given to DUs in the Netherlands. Thus, DUs seemed to enter the study because of HIV risk through injecting, not because of health problems.

With regard to demographics, LTM users were found to be more often Dutch and to have lived a longer time in Amsterdam than non–LTM users. This is in agreement with the intake rules of low-threshold programs, which discourage nonresident drug users from participating.

With regard to drug use, current daily injecting is similar for LTM and non–LTM users, as is borrowing of needles and syringes. LTM users inject mainly heroin less often and inject cocaine, either by itself or together with heroin, more often. A similar relation was found previously; Chaisson et al.[27] reports that methadone treatment reduces heroin injecting more strongly than it does cocaine injecting. Speedball injecting is associated with an increased HIV risk, both in the present sample and in New York City.[28]

To summarize, there were no indications that LTM users engaged in less HIV risk behavior than non–LTM users. Furthermore, when we looked at the effect of LTM use on HIV risk, while controlling for confounders, we found no indication of a protective effect of LTM use. The same was true in a separate analysis of incident HIV infections[19]: a protective effect of daily methadone use could not be identified.

These findings suggest that the low-threshold programs do not have an effect on HIV risk because risk behavior is not reduced. Another possibility is that a risk reduction takes place shortly after the start of methadone prescription but does not grow over the years. In contrast, in a prospective study among IDUs in US methadone maintenance programs,[7] the prevalence of injecting declined from 81% at the beginning of admission to 63% at the end of the period. After this prevalence declined gradually to 42% in the third year and 29% in the fourth year of participation. However, no such decline is apparent in our sample.

To conclude, our results do not support the view that long-term regular participation—as compared with short-term and/or irregular participation—in low-threshold methadone programs in Amsterdam is associated with less HIV risk–injecting behavior or with a decreased risk of HIV. The apparent failure of the programs to reduce such behavior may be due to the low average methadone dose level or to the (permitted) irregular attendance.[1,2,29] However, raising methadone doses to a level at which little or no illicit opiate use occurs, combined with measures to increase attendance, may lower the number of clients, given that not all DUs wish to stop illicit use. Such measures thus would counteract the possibility of providing medical and social care to a large group of DUs, many of whom are HIV seropositives.

It therefore seems necessary to judge carefully the relative merits of the low-threshold approach versus other kinds of approaches. In any case, we suggest that contacts with IDUs in the low-threshold programs should be used more intensively for HIV prevention efforts and that an increase in methadone dose level, together with measures to enforce regular attendance, may be required if these programs want to play a role in HIV prevention. □

Acknowledgments

This study was supported by the Netherlands Foundation for Preventive Medicine (grant no. 28–1258). The authors thank the nurses B. Frolich, B. Scheeringa-Troost, G. Stienstra, and J. Teeuwissen for interviewing and collecting blood samples; Dr. J. Goudsmit and M. Bakker for performing the laboratory tests; G.H.A. van Brussel and P. Drotman for critical comments on earlier drafts; H.J.A. van Haastrecht for critical comments and data management; and M. ter Pelle and T. Maruanaya for preparing the manuscript.

References

1. Cooper JR. Methadone treatment and acquired immunodeficiency syndrome. *JAMA*. 1989;262:1664–1668.
2. Dole VP. Methadone treatment and the acquired immunodeficiency syndrome epidemic. *JAMA*. 1989;262:1681–1682.
3. Novick DM, Joseph H, Croxson TS, et al.

Absence of antibody to human immunodeficiency virus in long-term, socially rehabilitated methadone maintenance patients. *Arch Intern Med.* 1990;150:97–99.
4. Zweben JE, Payte JT. Methadone maintenance in the treatment of opioid dependence, a current perspective. *West J Med.* 1990;152(special issue):588–599.
5. Zweben JE, Sorenson JL. Misunderstandings about methadone. *J Psychoactive Drugs.* 1988;20:275–281.
6. Dole VP. Methadone maintenance treatment for 25 000 heroin addicts. *JAMA.* 1971;215:1131–1134.
7. Ball JC, Lange WR, Myers CP, Friedman SR. Reducing the risk of AIDS through methadone maintenance treatment. *J Health Soc Behav.* 1988;29:214–226.
8. van Brussel GHA. Openbare gezondheidszorg en drugproblematiek. *T Soc Gezondheidszorg.* 1986;64:817–819.
9. Buning EC, Coutinho RA, van Brussel GHA, van Santen GW, van Zadelhoff AW. Preventing AIDS in drug addicts in Amsterdam. *Lancet.* 1986;1:1435. Letter to the Editor.
10. van Brussel GHA, Buning EC. Methadon per bus in Amsterdam. *T Alc Drugs.* 1979;5:135–137.
11. Buning EC. AIDS-related interventions among drug users in the Netherlands. Presented at the European Symposium on AIDS and Drug Abuse; September 25–28, 1989; Stockholm, Sweden.
12. Buning EC. *De GG & GD en het drugprobleem in cijfers, II, III, IV.* Amsterdam, The Netherlands: GG&GD; 1987, 1988, 1990.
13. Plomp HN, Reijneveld SA. *Methadon verstrekking aan druggebruikers in Amsterdam, analyse van gegevens van de Centrale Methadon Registratie over de jaren 1984 tot 1987.* Amsterdam, The Netherlands: VU utigeverij; 1991.
14. Driessen FMHM. *De methadonverstrekking in Amsterdam, 1981–1984, resultaten von de analyse van de gegevens van de Centrale Methadon Registratie.* Amsterdam, The Netherlands: Vrije Universiteit en GG&GD; 1987.
15. van den Hoek JAR, van Haastrecht HJA, Coutinho RA. Risk reduction among intravenous drug users in Amsterdam under the influence of AIDS. *Am J Public Health.* 1989;79:1355–1357.
16. van den Hoek JAR, Coutinho RA, van Haastrecht HJA, van Zadelhoff AW, Goudsmit J. Prevalence and risk factors of HIV infections among drug users and drug-using prostitutes in Amsterdam. *AIDS.* 1988;2:55–60.
17. van den Hoek JAR, van Haastrecht HJA, Scheeringa-Troost B, Goudsmit J, Coutinho RA. HIV infection and STD in drug-addicted prostitutes in Amsterdam: potential for heterosexual HIV transmission. *Genitourin Med.* 1989;65:146–150. Goudsmit J, Couthinho RA.
18. Miettinen O. Estimability and estimation in case-referent studies. *Am J Epidemiol.* 1976;103:226–235.
19. van Ameijden EJC, van den Hoek JAR, van Haastrecht HJA, Coutinho RA. The harm reduction approach and risk factors for HIV-seroconversion in injecting drug users in Amsterdam, the Netherlands. *Am J Epidemiol.* In press.
20. Novick D, Kreek MJ, Des Jarlais D, et al. Antibody to LAV, the putative agent of AIDS, in parenteral drug abusers and methadone-maintained patients: therapeutic, historical, and ethical aspects. National Institute on Drug Abuse Research Monograph 1986;67:318–320.
21. Watters JK, Newmeyer JA, Feldman HW, Biernacki P. Street-based AIDS prevention for intravenous drug users in San Francisco: prospects, options, and obstacles. In: *Community Epidemiology Work Group Proceedings,* Vol II. Rockville, Md: National Institute on Drug Abuse; 1986.
22. Hartel D, Selwyn PA, Schoenbaum EE, Klein RS, Friedland GH. Methadone maintenance treatment and reduced risk of AIDS and AIDS-specific mortality in intravenous drug users. Presented at the Fourth International Conference on AIDS; June 12–16, 1988; Stockholm, Sweden.
23. Blix O, Gronbladh BA. AIDS and IV heroin addicts: the preventive effect of methadone maintenance in Sweden. Presented at the Fourth International Conference on AIDS; June 12–16, 1988; Stockholm, Sweden.
24. Williams AB, Vranisan K, Gorter R, Brodie B, Meakin R, Moss A. Methadone maintenance, HIV-1 serostatus and race in injection drug users in San Francisco, California. Presented at the Sixth International Conference on AIDS; June 20–24, 1990; San Francisco, Calif.
25. Coutinho RA, Bos JM, Ruitenberg EJ. The epidemiology of LAV/HTLV-III infections in Europe. In: Staquet M, Hemmer R, Baert A, eds. *Clinical Aspects of AIDS and AIDS-Related Complex.* Oxford, England: Oxford University Press; 1986:7–14.
26. Reijneveld SA: *Methadonverstrekking in Amsterdam in 1989, jaaroverzicht van de Centrale Methadon Registratie.* Amsterdam, the Netherlands: GG&GD; 1990.
27. Chaisson RE, Bacchetti P, Osmond D, et al. Cocaine use and HIV infection in intravenous drug users in San Francisco. *JAMA.* 1989; 261:561–565.
28. Novick DM, Trigg HL, Des Jarlais DC, et al. Cocaine injection and ethnicity in parenteral drug users during the early years of the human immunodeficiency virus (HIV) epidemic in New York City. *J Med Virol.* 1989;29:181–185.
29. Dole VP. Implications of methadone maintenance for theories of narcotic addiction. *JAMA.* 1988;260:3025–3029.

The Potential Role of Custody Facilities in Controlling Sexually Transmitted Diseases

Deborah Cohen, MD, MPH, Richard Scribner, MD, MPH, John Clark, MD, MPH, and David Cory

ABSTRACT

Background. **The National Commission on Correctional Health Care recommends that medical screening, including tests for syphilis and other sexually transmitted diseases, be performed by the 14th day after initial booking procedures. Because the average length of stay in detention is usually less than 14 days, most detainees are not screened or treated for asymptomatic communicable diseases.**

Methods. **In order to determine the incidence and prevalence of syphilis among those booked through a large county jail, a rapid screening procedure to test newly incarcerated inmates and treat those infected was implemented over a 3-week period.**

Results. **Among 6309 detainees eligible for testing during the 3-week period of the study, 129 (2%) were diagnosed with syphilis. During the preceding month when only those reporting symptoms or high-risk behaviors were screened, 20 cases were diagnosed. The sensitivity and predictive value of self-reported risk factors for syphilis infection were low. Among the first 110 diagnosed cases, 17 (15.5%) were released within 3 days and 86 (78%) were released within 6 days.**

Conclusion. **If screening and treatment for syphilis are delayed until the 14th day after booking, the majority of infected inmates will be released prior to evaluation and treatment. Rapid screening and treatment for selected sexually transmitted diseases are likely to be an effective public health measure when implemented in custody settings. (*Am J Public Health.* 1992;82:552–556)**

Introduction

Several studies have documented the high prevalence of sexually transmitted diseases (STDs) among persons in custody facilities.[1–3] Screening for STDs among both adult and juvenile offenders is mandatory in some states and is recommended by the American Medical Association.[4] The National Commission on Correctional Health Care recommends that medical screening or review of medical screening be performed on or before the 14th day after initial booking procedures.[5] Because the average length of stay in detention is usually less than 14 days,[6] most persons arrested are not screened or treated for asymptomatic communicable diseases. Even when screening is performed at the time of booking, results may not be available until after the inmate has been released. Subsequently, those at risk may not be located and treated. Considering the rapid turnover of inmates in the custody setting, rapid screening, diagnosis, and treatment may significantly reduce STD morbidity.

Background

Beginning in 1986, Men's Central Jail (MCJ), the main custody facility for men in Los Angeles County (LAC), offered a serologic test for syphilis (STS) on admission only to persons who self-identified any of several risk factors, i.e., homosexuality, history of STD, intravenous drug use, and current symptoms of STD. This policy was instituted primarily to reduce costs. It was assumed that the yield of true positives would be greater through a self-identification process. In addition, all inmates would receive a full health appraisal after 2 weeks of residence.

In 1988, cases of syphilis reported from custody facilities comprised more than 10% of the county's early syphilis morbidity, even though fewer than 10% of the persons booked through custody facilities were tested.[7] The LAC jails comprise the largest custody facilities in the United States, booking approximately 250 000 persons annually and having about 25 000 residents on any given day.[6]

In order to determine the incidence and prevalence of syphilis among those booked through MCJ and to establish a risk profile of infected inmates that might be more objective than self-identification, a 3-week project to test newly incarcerated inmates was implemented. Because inmates who are released are difficult to locate for further evaluation and treatment, the protocol included immediate STS results and treatment prior to release.

Methods

Between March 29, 1989, and April 19, 1989, the Sexually Transmitted Disease Program of the Los Angeles County Department of Health Services provided a "stat" rapid plasmin reagin (RPR) test to all inmates booked at MCJ between 6 PM and 4 AM, Monday through Friday, when

Deborah Cohen, Richard Scribner, and David Cory are with the Los Angeles County Department of Health Services. Deborah Cohen and Richard Scribner also are associated with the University of Southern California. John Clark is with the Los Angeles County Sheriff's Department.

Requests for reprints should be sent to Deborah Cohen, MD, MPH, Department of Family Medicine, USC, 1420 San Pablo B-205, Los Angeles, CA 90033.

This paper was submitted to the journal September 14, 1990, and accepted with revisions April 4, 1991.

the majority of men held for at least 24 hours are processed.

Because syphilis screening in custody facilities is mandated by law, all inmates were asked to submit to phlebotomy. Inmates who refused to have their blood drawn were separated from the others and referred for an interview and exam by project staff. To obtain a random subsample of nonreactor inmates for comparison, the third inmate whose RPR was nonreactive out of a batch of 10 to 15 specimens tested was also evaluated. A venereal disease research laboratory (VDRL) test and a confirmatory microhemagglutination assay for *Treponema pallidum* (MHA-TP) was performed on all inmates with reactive RPR as well on nonreactor controls.

Demographic data including age, race, and criminal charges were obtained from custody records. Public health nurses administered a questionnaire assessing sexual behavior, history of sexually transmitted diseases, and drug use. Physicians examined inmates with a history of symptoms or a reactive RPR. Inmates with reactive tests were interviewed for contacts, and named contacts were referred for field follow-up by health department investigators. Record searches through the Los Angeles County Venereal Disease Record System (VDRS) were performed for all persons with confirmed reactive serologies to compare prior titers and verify a history of prior diagnosis and treatment.

An inmate was classified as having primary syphilis if the physicians documented a typical chancre, regardless of RPR result. A diagnosis of secondary syphilis was made based upon typical lesions and confirmed reactive treponemal tests. Early latent syphilis was diagnosed if the patient had a reactive serology confirmed by a reactive MHA-TP, a history of typical lesions or negative serology in the previous 12 months or sexual contact with another case of early syphilis, and no documentation or history of treatment. Any asymptomatic inmate who had a confirmed reactive RPR, had no history of treatment, or did not fall into the previous categories was diagnosed with late latent syphilis. A case was defined as the presence of untreated syphilis, regardless of stage.

Inmates with reactive RPR, no history of treatment, or physical findings consistent with syphilis were offered treatment with benzathine penicillin G or with doxycycline in case of penicillin allergy. Those with a history of prior treatment or a weakly reactive RPR and no physical findings were treated at the discretion of the project physicians.

All data were analyzed using the Statistical Analysis System (SAS) and Epi-Info, which uses the Cornfield method for determining odds ratios (ORs) and confidence intervals (CIs).

TABLE 1—Race-Specific Rates of Infectious Syphilis among Men Booked at Men's Central Jail, Los Angeles County, March 26–April 19, 1989

Race/ Ethnicity	N	Infectious Syphilis	Early Latent Syphilis	Late Latent Syphilis	Total Syphilis
Whites	1292	0 (0)	0 (0)	1 (77)	1 (77)
Blacks	2135	21 (984)	22 (1030)	24 (1124)	67 (2998)
Hispanics	2632	11 (418)	21 (798)	27 (1026)	59 (2242)
Asians	84	0 (0)	0 (0)	0 (0)	0 (0)
Other/Unknown	166	0	0	2	2
Total	6309	32 (507)	43 (682)	54 (856)	129 (2045)

Note. Data are number of cases (rate per 100 000).

Results

Population Description

During the 3-week project period, 12 147 men were booked through MCJ. The study population comprised 6309 men who were booked for at least 24 hours between 6 PM and 4 AM and offered an STS. Of these, 95 refused phlebotomy, leaving 6214 men tested. The remaining 5867 inmates not in the study sample were either booked during daytime hours or were immediately released on their own recognizance. No data are available for this group.

The study population was predominantly Black and Latino (75.6%). The average age was 30. Forty-eight percent of the inmates were charged with alcohol- or drug-related offenses, 22.2% with theft-related charges, 9.6% with crimes of violence, 7.2% with traffic violations, and 1.5% with crimes related to sexual behavior. The majority of the study group said they lived in LAC.

Rates of Syphilis

Of the men screened for syphilis, the RPR was reactive in 308 (4.9%). Physicians reported active lesions in 32 persons, one of whom had a nonreactive RPR and was diagnosed because he had been part of the interviewed nonreactor group and had symptoms. Another case of primary syphilis was found in an inmate who had refused the blood test but was examined. Untreated syphilis was identified in 129 men for a rate of 2045/100 000. The rate of infectious syphilis was 507/100 000. (The county-wide rate was 45.1/100 000 in 1988.) Early latent syphilis was diagnosed in 43 men for a rate of 682/100 000. The race-specific rate for early syphilis was highest among Blacks (2014/100 000). Fifty-four men had latent syphilis infection of unknown duration which we classified as late latent syphilis (Table 1). Of the 308 men with reactive RPRs, 75 (24%) had a negative MHA-TP and 84 (27%) were patients who had a history of prior treatment for syphilis.

Time of Release of Cases

Among the first 110 diagnosed cases, 17 (15.5%) were released within 3 days and 86 (78%) were released within 6 days. Twenty-one of the 233 inmates with a reactive MHA-TP were not available for further evaluation and so a diagnosis could not be made for them.

Risk Factors

Because of time and security constraints, inmates with reactive RPRs were preferentially evaluated over nonreactors. Among reactors, 278/308 (90%) were evaluated, among refusers 33/95 (35%), and among nonreactors, 372/531 (70%).

Ninety-eight percent of all cases were either Black or Latino and their risk of syphilis was, respectively, 41.7 and 29.6 times that of Whites. Cases were more likely than noncases to have a history of syphilis, gonorrhea, or any STD and were more likely than controls to have had three or more sex partners in the previous 90 days (OR=3.46, 95% CI=1.48–8.26). Only three cases (2.3%) and four nonreactors (1.1%) admitted to ever having sex with other men. These seven men said they also had sex with women. Cases were more likely to be between the ages of 31 and 40 (Table 2). Inmates with syphilis were less likely to be married and more likely to say they had less education than nonreactors.

TABLE 2—Characteristics of Study Sample

Characteristic	Cases n (%)	Noncases[a] n (%)	OR (95% CI)[b]
Race			
White	1 (0.8)	1289 (20.9)	1.00
Black	67 (51.9)	2070 (33.5)	41.72 (7.21–1,673.46)
Hispanic	59 (45.7)	2573 (41.6)	29.56 (5.09–1,187.89)
Asian	0 (0)	84 (1.4)	
Other/unknown	2 (1.6)	164 (2.7)	
Total	129	6180	
Age			
15–20	0 (0)	311 (5.1)	
21–25	23 (18.3)	1515 (24.9)	1.00
26–30	31 (24.6)	1547 (25.4)	1.32 (0.74–2.35)
31–35	35 (27.8)	1248 (20.5)	1.85 (1.05–3.25)
36–40	21 (16.7)	660 (10.8)	2.10 (1.11–3.96)
41–45	9 (7.1)	392 (6.4)	1.51 (0.64–3.46)
>45	7 (5.6)	414 (6.8)	1.11 (0.43–2.75)
Missing	3	93	
Mean age	31.7	30.4	
Booking charge			
Alcohol related	13 (10.7)	1268 (20.9)	0.65 (0.24–1.80)
Drug related	39 (32.2)	1684 (27.8)	1.46 (0.62–3.60)
Sex related crimes	3 (2.5)	91 (1.5)	2.08 (0.42–9.12)
Theft/burglary	38 (31.4)	1334 (22.0)	1.79 (0.76–4.43)
Violence/firearms	11 (9.1)	580 (9.6)	1.19 (0.43–3.44)
Other	10 (8.3)	663 (10.9)	0.95 (0.33–2.78)
Traffic violations	7 (5.8)	441 (7.3)	1.00
Missing	8	119	

[a]Noncases are all inmates with no diagnosis of syphilis.
[b]OR=odds ratio; CI=confidence interval.

There was no association between booking charge and syphilis infection.

Only 3 of the 32 inmates with infectious syphilis identified themselves as having symptoms at the time of entry into the study. In a confidential setting, an additional 15 men said they had a genital "sore" or ulcer. Upon examination, 5 of these men did not have syphilitic lesions.

Cases were more likely than nonreactors to say they regularly used smokable crack cocaine (OR=2.07, 95% CI=1.16–3.69). No other drug was associated with syphilis cases more than with noncases. The sensitivity and positive predictive value of the hypothesized risk factors obtained in a confidential setting were low (Table 3).

Refusers

Compared to cases, the 33 refusers were more likely to be White (21% vs 1% of cases), to be married (27% vs 13%), and to say they had had a blood test for syphilis in the previous 12 months (39% vs 15%).

False-Positive Results

Among 75 inmates who had false positive RPR results, 44.7% reported regular heroin use and 70% said they had tried heroin at least once. These inmates were older than cases (average age=36.5 years) and more of them were White (24% vs 1%). They were more likely to say they had a history of syphilis or "bad blood" (23% vs 8%) and more likely to have been booked for alcohol-related charges (35% vs 11%).

Contact Interviewing and Tracing

Of the 308 patients with reactive, qualitative RPRs, 187 were interviewed by investigators. During the interview process, 29 said they had been treated previously and another 29 were subsequently found not to be infected based on a negative MHA-TP. Interviewers elicited 93 new contacts from 60 interviewed inmates with early syphilis. Of these, 31 (33%) were located during field follow-up. Eight of the contacts had early syphilis. Field investigations were initiated for acquaintances of reactors who were pregnant. Out of eight pregnant women named, the only one found infected was a prostitute with secondary syphilis. Preventive treatment was given to 19 contacts. One contact had been previously treated and one refused evaluation. The overall contact index (number of contacts elicited per case) was 1.6, and the disease intervention index (percentage of cases where at least one contact was treated[8]) was 0.32.

Discussion

This study documents a 2% rate of syphilis among inmates booked through Los Angeles County's Men's Central Jail and the feasibility and effectiveness of integrating rapid screening and treatment into the booking process. In the month prior to the screening project, 20 cases of syphilis were reported by the jail compared to 129 cases during the 3-week project.

Usefulness of Risk Factor Identification

The attempt to identify factors that could be used to select out the highest risk inmates with a high degree of sensitivity and specificity was unsuccessful. In the jail setting with criminal charges pending, inmates are unlikely to bring special notice to themselves or to be forthright about behaviors that are illegal. Although the study verified that a history of gonorrhea or syphilis and the presence of symptoms elicited in a confidential setting were associated with syphilis, these risk factors had low sensitivity and positive predictive value. Other risk factors such as homosexuality and heroin use did not appear as risk factors, probably because so few admitted to these behaviors.

Risk factors identified through the case-reference methodology included the use of crack cocaine, three or more sex partners in the previous 90 days, history of syphilis or gonorrhea, and less than 10th-grade education. However, if a screening program depended on obtaining a history of drug use or high-risk sexual behavior in a confidential setting, approximately half of the infected cases would be missed. Given the rapidity with which men are booked, the crowded conditions, and the lack of confidential settings in the county jail, individual history taking for all inmates is not feasible.

Race (i.e., being Black and Latino) was the risk factor with the highest sensitivity (98%) for syphilis, reflecting the county-wide trend in Los Angeles: compared to Whites, Blacks were at a 30-fold and Hispanics at an 8-fold greater risk of being reported with infectious syphilis.[9]

Limitations

We could not interview all members in the reference group; however, there was no difference between the reference group and the entire population in the distribution of race, age, and booking charge.

We believe there was no systematic bias in selecting the reference group and our findings are likely to be generalizable to the entire population of men booked at night for criminal offenses in Los Angeles.

Recommendations

The criminal justice setting may be one of the few public institutions in which there is an opportunity to break chains of infection among high-risk populations. Rapid screening and treatment of STD in custody facilities should be considered public health control measures in localities with a high incidence of STDs. When the development of antibiotic resistance is not a concern and the incidence of serious adverse reactions is minimal, all inmates with laboratory evidence suggestive of disease should be empirically treated. In our study we treated very few inmates whose qualitative RPR was weakly reactive. Eleven inmates subsequently had titers ≥1:16 and reactive MHA-TP tests but could not be found for treatment.

Contact tracing was generally unproductive. The achieved disease intervention index (0.32) compares poorly to expected national standards of 0.60. Results might have been higher if the interviewer had been the person assigned to field follow-up. Given the morbidity and mortality of congenital syphilis, intensive efforts to locate only contacts or suspects who are pregnant may be useful.

While we have not explored the feasibility of screening and treatment for other STDs such as gonorrhea and chlamydia, the prevalence of these STDs is likely to be higher[1] and the cost of screening men with the leukocyte esterase dipstick substantially lower than with the RPR.[10] Syphilis, gonorrhea, and chlamydia have been found to be potential risk factors for human immunodeficiency virus (HIV) disease.[11–13] A substantial reduction in morbidity might occur, not only for transmission of HIV, but also for other sequelae such as infertility, pelvic inflammatory disease, and ectopic pregnancies.[14]

Few public resources are devoted to providing preventive medical care for persons charged with criminal offenses; yet, the public health impact of treating a large infected population is likely to be enormous in decreasing the chance that many other individuals in the community will become infected. Population-based screening of groups characterized by the societal conditions associated with high-risk behavior and lack of access to health care, such as socioeconomic status, race, and place of residence, is likely to be more effective in limiting STDs than control programs that require individuals to identify themselves to be at risk and to actively seek screening and treatment. □

TABLE 3—Factors Related to Syphilis Infection

Factors	Cases n (%)[a]	Reference Group n (%)	OR (95% CI)[b]	PPV[c]
Marital status				
Single	77 (67)	210 (62)	1.00	0.02
Married	15 (13)	89 (26)	.46 (0.24–0.87)	0.01
Other (divorced separated, widowed)	23 (20)	38 (11)	1.65 (0.89–3.06)	0.03
Education				
<9 years	40 (31)	73 (19)	2.49 (1.16–5.41)	0.03
10–12 years	75 (58)	240 (65)	1.42 (0.71–2.88)	0.02
>12 years	13 (10)	59 (16)	1.00	0.01
Sexual partners (in last 90 days)				
0	10 (8)	49 (14)	1.00	0.01
1	46 (38)	183 (53)	1.23 (0.55–2.81)	0.01
2	23 (19)	53 (15)	2.13 (0.86–5.36)	0.02
≥3	41 (34)	58 (17)	3.46 (1.48–8.26)	0.04
Sexual preference				
Sex with men and women	3 (2)	4 (1)	3.02 (0.63–14.62)	0.05
Sex with women	121 (98)	366 (99)	1.00	0.02
STD history				
Syphilis	10 (8)	23 (6)	2.22 (1.04–4.74)	0.03
Gonorrhea	54 (45)	109 (31)	1.86 (1.19–2.92)	0.03
Other STD	9 (7)	20 (6)	1.78 (0.73–4.26)	0.02
Any STD in the past	60 (49)	133 (37)	1.67 (1.08–2.59)	0.03
No STD in the past	62 (51)	230 (63)	1.00	0.02
Smokable cocaine use				
Regular use	22 (17)	49 (9)	2.07 (1.16–3.69)	0.04
No regular use	106 (83)	486 (91)	1.00	0.02
Heroin use				
Ever used	20 (16)	90 (24)	0.32 (0.18–0.55)	0.01
Never used	108 (84)	282 (76)	1.00	0.02

[a]Sensitivity
[b]OR=odds ratio (calculated using Cornfield method); CI=confidence interval.
[c]PPV=positive predictive value.

Acknowledgments

The authors wish to thank Drs. Laurene Mascola, Clyde Dent, Susan Turner, and Ray Palmer for their valuable assistance.

References

1. Moran JS, Peterman T. Sexually transmitted diseases in prisons and jails. *Prison J.* 1989;69(2):1–6.
2. Alexander-Rodriguez T, Vermund SH. Gonorrhea and syphilis in incarcerated urban adolescents: prevalence and physical signs. *Pediatrics.* 1987;80:561–564.
3. Martin JW, Much DH. Sexually transmitted diseases in prison women. *Pa Med.* April, 1988;91:40–42.
4. American Medical Association. *Standards for Health Services in Jails.* Chicago, Ill: American Medical Association; 1979.
5. National Commission on Correctional Health Care. *Standards for Health Services in Jails.* Chicago, Ill: National Commission on Correctional Health Care; 1987.
6. *Sheriff's Annual Fiscal Summary, July 1989–June 1990.* Los Angeles County Sheriff's Department; 1990.
7. *Venereal Disease Record System, S-50, Sexually Transmitted Disease Program.* Public Health Programs and Services, Los Angeles County Department of Health Services; 1988.
8. Centers for Disease Control. *Guidelines for STD Control, Program Operations.* Atlanta, Ga: Centers for Disease Control; 1985. US Dept of Health and Human Services.
9. *Communicable Disease Morbidity Report, 1988.* Los Angeles, Calif: Los Angeles County Department of Health Services.
10. Sadof MD, Woods ER, Emans J. Dipstick leukocyte esterase activity in first catch urine specimens: a useful screening test for detecting sexually transmitted disease in the adolescent male. *JAMA.* 1987;258: 1932–1934.
11. Cameron DW, Simonsen JN, D'Costa LJ, et al. Female to male transmission of human immunodeficiency virus type 1: risk factors for seroconversion in men. *Lancet.* 1989;2:403–407.

12. Latif AS, Katzenstein DA, Bassett MT, Houston S, Emmanuel JC, Marowa E. Genital ulcers and transmission of HIV among couples in Zimbabwe. *AIDS*. 1989;3:518–523.

13. Pepin J, Plummer FA, Brunham RC, Piot P, Cameron W, Ronald AR. The interaction of HIV infection and other sexually transmitted diseases: an opportunity for intervention. *AIDS*. 1989;3:3–9.

14. Moore DE, Cates W. Sexually transmitted diseases and infertility. In: Holmes KK, Mardh PA, Sparling PF, et al., eds. *Sexually Transmitted Diseases*. New York, NY: McGraw-Hill; 1990;63:763–770.

Workshop on Statistics & Computing in Disease Clustering to Be Held July 23 & 24, 1992, in Stony Brook, NY

Applied Biomathematics, with funding from the Electric Power Research Institute and the National Cancer Institute, is conducting a 2-day workshop on statistical methods and computer resources for disease clustering. This field is changing rapidly and novel approaches for analyzing clusters are frequently proposed. The aim of this workshop is the active involvement of the participants in discussion of the following areas:

Scientific Merit How should clustering methods be selected, and what can clustering studies offer?

Cluster Allegations How should we deal with post hoc data analyses?

Cluster Patterns How can patterns in space and time be characterized?

Methods What is the power of clustering techniques under different space–time arrangements of cases?

Applications What research needs are revealed by past analyses?

GIS How can geographic information systems and geographic statistics be used to identify links between environment and health?

Resources What databases and microcomputer software are currently available for cluster investigations?

Contributed papers are solicited on these or related topics. The program committee consists of Dr. Geoffrey Jacquez, Applied Biomathematics (chair); Dr. Roger Grimson, Department of Preventive Medicine, SUNY at Stony Brook; Dr. Leeka Kheifets, Electric Power Research Institute; and Dr. Daniel Wartenberg, Robert Woods Johnson Medical School.

This workshop will bring together those who develop statistical and computational methods with those who use them. Workshop information may be obtained from Dr. Jeffrey A. Millstein, Applied Biomathematics, 100 North Country Road, Setauket, New York 11733. Tel: 516/751-4350; Fax: 516/751-3435.

Occupational Injuries among Minors Doing Farm Work in Washington State: 1986 to 1989

Nicholas J. Heyer, PhD, Gary Franklin, MD, MPH, Frederick P. Rivara, MD, MPH, Paul Parker, JD, and Joanna A. Haug, MS

ABSTRACT

Background. There is growing evidence that many children are injured while engaged in agricultural work. However, little specific information on farm work–related injuries among minors is available, probably because employment or workers' compensation data for children are hard to obtain.

Methods. Workers' compensation data were used to evaluate occupational injuries among children in Washington State from 1986 through 1989. The frequency and severity of injuries among minors doing farm work were compared with the distributions of injuries among minors working in food service and all other occupations by year of injury, age of injury, and month and hour of injury.

Results. A total of 16 481 claims filed by children under age 18 were evaluated. Although farm workers accounted for only 7% of all claims, they made up 36% of claims filed by children under age 14, and 17% of claims filed by children aged 14 or 15. Injuries classified as serious accounted for 26% of farm worker claims compared with only 16% of all claims filed by children.

Conclusions. Although injury rates could not be developed owing to the lack of denominator data, this study demonstrates that farm work is dangerous for young children. (*Am J Public Health.* 1992;82:557–560)

Introduction

Agriculture is one of the most hazardous industries in the United States, sharing with mining and construction the highest rates of occupational mortality.[1] Yet young children frequently work on farms and are subject to this risk. It has been estimated that 14% to 24% of all fatal injuries on farms (occupational and nonoccupational) occur among children 15 years of age or younger,[1] and that persons 19 years of age or younger are involved in 25 000 nonfatal traumatic events and 300 injury deaths on farms annually in the United States.[2]

Although numerous reports have looked at farm injuries among children in various states,[1,3–11] few have focused specifically on occupational (i.e., work-related) farm injuries among children. Two hospital-based reports on farm injuries among minors tried to focus on work-related injuries by excluding drownings, shootings, and poisonings. These reports demonstrated high levels of injury severity (largely due to involvement with machinery).[4,11] A study of fatal injuries involving tractors reported that 29% occurred among children aged 19 or younger.[7]

Methods

The Washington State workers' compensation program (administered by the state's Department of Labor and Industries) covers all employers with one or more full- or part-time employees. Major exemptions include employees covered by some other compensation program (e.g., federal, police and fire fighters') or, for agricultural employment, family members of the farm owner who are minors (under 18 years of age). Workers' compensation coverage may be through self-insurance or the State Fund. Two thirds of nonfederally employed workers in Washington State are covered by the State Fund.[12]

All claims for workers' compensation filed by minors through the State Fund were reviewed for the 4 years, 1986 through 1989. Claims filed for either injury or occupational disease were included because it is often difficult to distinguish between the two. However, disease claims represent less than 4% of all claims.[13]

The following data elements were used for our analysis:

1. US Bureau of Census Occupation Codes,[14] to define farm workers (code 479) and food service workers (codes 435 through 444). All claims not in one of these two categories were coded as "all others." Farmers (code 473), farm managers (code 476), and farm supervisors (code 477) were not included among farm workers, as they do not represent the farm employment of interest.
2. Age at time of injury, categorized into three groups: 13 years of age

Nicholas Heyer and Joanna Haug are with the Occupational Epidemiology and Health Outcomes Program of the University of Washington's Department of Environmental Health in Seattle, and Frederick Rivara is with its School of Medicine. Paul Parker is with the State of Washington's Department of Labor and Industries. Gary Franklin is affiliated with all three institutions.

Requests for reprints should be sent to Nicholas Heyer, PhD, Department of Environmental Health, SC-34, University of Washington, Seattle, WA 98195.

This paper was submitted to the Journal December 4, 1990, and accepted with revisions September 20, 1991.

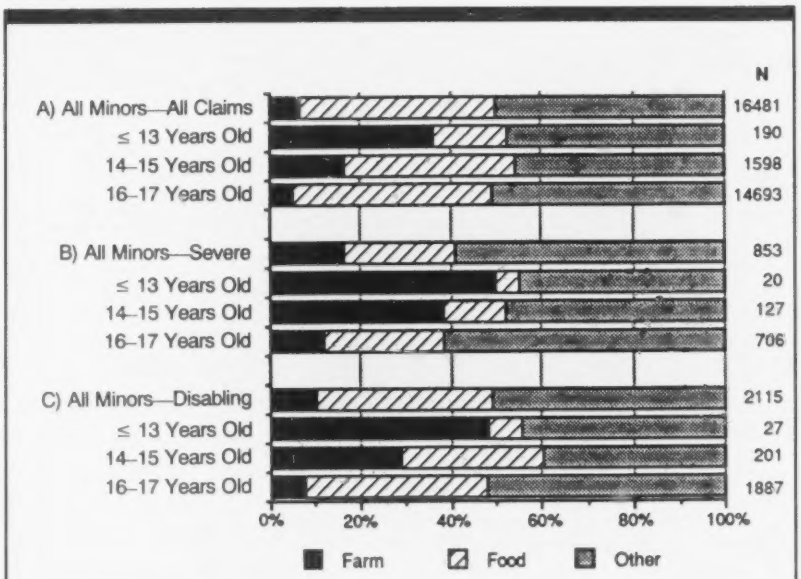


FIGURE 1—Distribution of occupations within age groups for all claims, severe claims, and disabling claims.

or younger, 14 or 15 years of age, and 16 or 17 years of age.

3. Year of injury.
4. Month of injury, coded to season: winter (December, January, and February), spring (March, April, and May), summer (June, July, and August), and fall (September, October, and November).
5. Hour of injury, coded to AM (after midnight through noon) or PM (after noon and through midnight).
6. Serious injuries, defined using two independent, surrogate measures: nature of injury and time loss payment (made by the State Fund after the loss of 4 or more days of work). Nature of injury was recorded using the American National Standards Institute Z-16 codes.[15] Only injuries with code 100 (amputations), code 110 (asphyxiation or drowning), code 140 (concussion), code 210 (fracture), or code 400 (multiple injuries) were considered to be "severe." All injuries with time loss payments were considered to be "disabling."

Specifically, the workers' compensation records were used to examine the frequency and severity of claims among minors employed as farm workers. Although rates were not estimated because total employment data are not maintained for minors, the distribution of claims among minors in farm worker, food service, and all other occupations was evaluated, as the distribution of claims among injuries classified severe, disabling, or both. The relationships between injuries, season, and time of day of injury were also evaluated because of the nature of farm employment and its potential interaction with school performance.

Results

All Claims

A total of 916 475 claims were filed with the Department of Labor and Industries during this 4-year period, with 16 481 (1.8%) filed by minors. Of the claims filed by minors, 190 (1.2%) were filed by children aged 13 or younger, 1598 (9.7%) by children aged 14 or 15, and 14 693 (89.2%) by children aged 16 or 17.

Figure 1 shows the distribution of occupations for all claims by age group. Although farm workers represent less than 7% of all claims filed by minors, they account for 36% of claims filed by children aged 13 or younger and for 17% of claims filed by children aged 14 or 15. By comparison, food workers represent 43% of all claims filed by minors, of which only 16% were filed by children in the youngest age group and 38% by children aged 14 or 15.

Serious Injuries

A total of 2617 (15.9%) claims filed by minors were for serious injuries (severe, disabling, or both). Among farm worker claims, 26% were for serious injuries. The relative frequencies of claims for severe and disabling injuries, severe only injuries, and disabling only injuries were all higher among farm workers than among food service workers or all other workers (Figure 2). The differences were greatest for the youngest age groups. Among farm workers, serious injuries accounted for 25% of all claims filed by children aged 13 years or younger and for 30% of all claims filed by children aged 14 or 15. By comparison, the percentages for food service workers were 12% and 6.5%, respectively.

Farm work contributes substantially to the total number of severe or disabling injuries among the two youngest age groups. Among workers aged 13 or younger, farm work accounted for 50% of all severe injury claims and for 48% of all disabling injury claims, compared with 5.0% and 7.4%, respectively, among food service workers. Among workers aged 14 or 15, farm workers accounted for 39% and 29% of severe and disabling injury claims, respectively, compared with 13% and 31% for food service workers (Figure 1 b and c).

Variation by Year and Season

There is no consistent pattern of variation in claims for minors by year and occupation. However, among food service workers, the two youngest age categories have almost doubled their proportion of claims, from 6.4% in 1986 to 12.4% in 1989.

There is, however, considerable variation by season for all occupational categories, especially for farm workers (Figure 3). The summer quarter accounts for 68% of all farm worker claims, compared with 38% of all food service worker claims and 52% of all other claims. The summer quarter also represents a disproportionately high percentage of claims among the two youngest categories of children (13 years of age or younger, or 14 or 15 years of age), who together account for 36% of all farm worker claims, 11% of all food service worker claims, and 13% of all other occupational claims.

Variation by Time of Day

Figure 4 shows the distribution of claims by time of day (AM/PM) within season for the three occupational categories. Farm workers filed claims for injuries occurring in the morning 58.9% of the time, compared with only 21.3% for food service workers and 36.2% for all other occupa-

tional claimants. During the three seasons when school is in session (i.e., excluding summer), the percentage of farm worker claims filed for morning injuries ranged from 32.4% to 56.0%, compared with a range of 18.8% to 19.6% for food service workers and of 26.0% to 28.4% for all other occupational claimants.

Discussion

The descriptive information obtained from Washington State's workers' compensation data about farm injuries among minors must be viewed as preliminary, given that no rates could be developed owing to the lack of denominator data. However, the data are consistent with previous reports in demonstrating that injuries related to farm work tend to have high levels of severity and that a substantial number of these injuries is occurring among young children.[4,7,11]

The distribution of all injuries, and of serious injuries by age, was compared for minors in farm work, in food service, and in all other occupational categories. The proportion of total claims filed among minors aged 13 or younger and among minors aged 14 or 15 is greater for farm workers than for food service workers and all other workers. This could be due to either a greater percentage of young minors being employed as farm workers or a greater percentage of those young minors employed as farm workers being injured, or both.

Furthermore, the proportions of severe injury claims, disabling injury claims, and both severe and disabling injury claims are all much higher among farm workers in all age categories than among food service workers and all other workers. The proportion of serious injury claims is particularly high among farm workers aged 14 or 15 (more than 30% of all claims). The proportion of serious claims among minors engaged in farm work is twice that among those employed in food service or all other employment. Unless minors employed in food service and other nonfarm employment file twice as many claims overall as those employed in farm work, the unmeasured, underlying rate of serious claims among minors employed in farm work will be greater.

Independent of any assumptions, it is clear that farm work accounts for a considerable proportion of all claims, especially serious injury claims among children aged 13 or younger and among children aged 14 or 15. Farm workers account for 49% of serious injury claims (severe, disabling, or both) among children aged 13 years or younger and 30% of such claims for children aged 14 or 15.

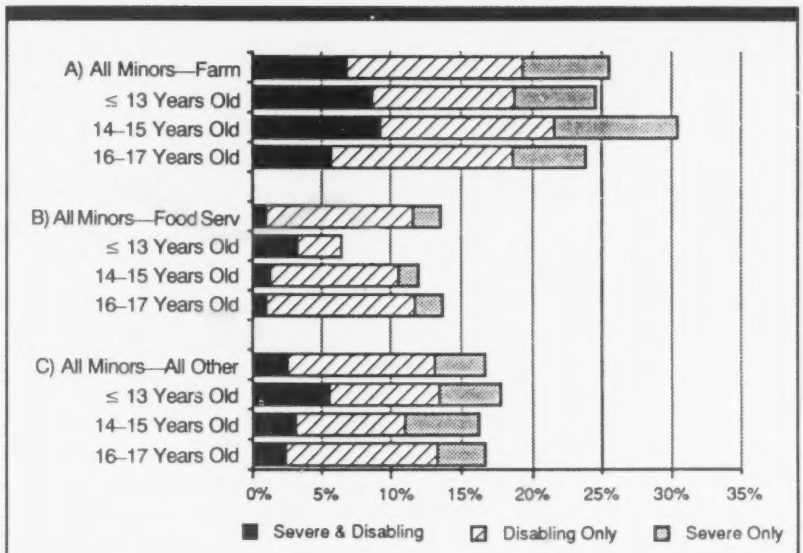


FIGURE 2—Percentage of severe and/or disabling claims within occupation and age groups.

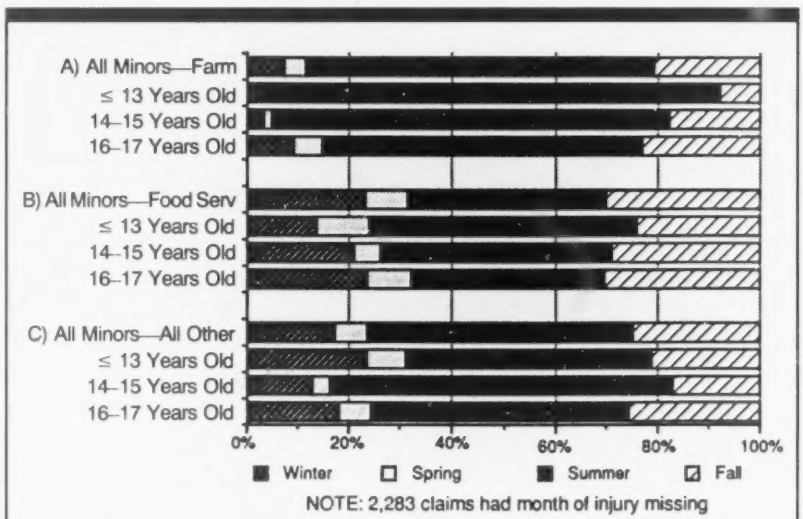


FIGURE 3—Distribution of claims by season within occupation and age groups.

Although no temporal trends for injury claims among farm worker minors could be identified, the proportion of claims among the two youngest categories of minors doing food service work nearly doubled between 1986 and 1989. This may represent that there has been a shift to younger employees or that more hazardous work is being performed by the youngest employees in the food service industry. This trend should be monitored.

The diurnal distribution of claims was also reviewed. A much greater percentage of farm worker injuries occurred in the morning during the winter and fall months when compared with injuries for food service employment or all other employment (Figure 4). This high percentage of morning claims suggests that farm worker minors (1) are working long hours before going to school, (2) are at particularly high risk during preschool working hours, (3) are working during morning school hours, or (4) are engaged in some combination of the above. Further investigation in regard to these factors and their possible interaction with the perceived high rate of school dropouts or

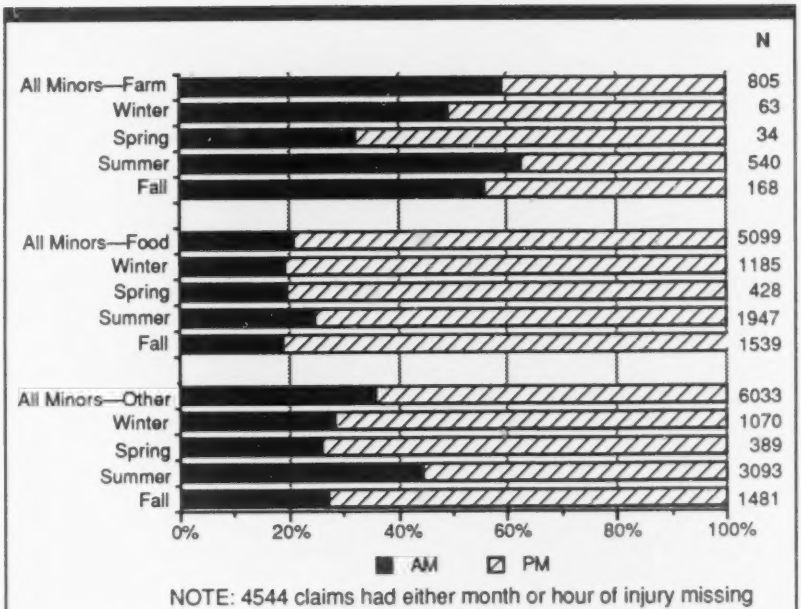


FIGURE 4—Distribution of claims by time of day and season within occupational group.

failure among farm worker children[16] requires further critical appraisal.

Concern about the potentially negative impact of farm work on the health and school performance of minors is not new. The data presented here were generated in support of the rule-making process that is currently ongoing in Washington State concerning employment of minors. Rules recently adopted by the Department of Labor and Industries[17] exceed federal standards in many respects. The state will prohibit minors under age 14 (who are not members of the farm owner's family) from doing farm work except for summer hand harvesting and cultivation of four crops (berries, flower bulbs, cucumbers, and spinach). The regulations will limit the number of hours minors can work per day and per week, and they will prohibit the work minors can do during certain night and early morning hours.

Using workers' compensation data has several drawbacks. First, injuries are likely to be underreported because only those injuries resulting in the filing of a claim are recorded. The underreporting is likely to be more severe for farm injuries than for other employment categories (especially food service, in which child labor is well established and occurs largely in well-supervised, national chains). Thus, the data presented here probably underestimate the farm injury problem. Second, the unavailability of data to directly evaluate rates of injuries among minors is of concern. The fact that worker hours are not reported for minors in general, and for farm worker minors in particular, is a major methodological problem in addressing the issues raised in this paper. □

Acknowledgments

This work was funded by the Accident and Medical Aid Funds of the Washington Department of Labor and Industries. These research monies are targeted to reducing the incidence of and disability related to occupational injuries and illnesses.

References

1. Field WE, Purschwitz MA. Cost of farm and rural injuries. *Public Health Rep.* 1987;102(6):642–644.
2. Rivara FP. Fatal and nonfatal farm injuries to children and adolescents in the United States. *Pediatrics.* 1985;76(4):567–573.
3. Rivara, FP. Farm-tractor associated deaths: Georgia. *MMWR.* 1983;32(37): 481–482.
4. Cogbill TH, Busch HM, Stiers GR. Farm accidents in children. *Pediatrics.* 1985;76(4):562–566.
5. Davis JB, Howell CG, Parrish RA. Childhood farm injury: the role of the physician in prevention. *Am Surg.* 1988;54:192–194.
6. Jackson FC. Farm and ranch injuries in West Texas. *Tex Med.* 1983;79:51–54.
7. Karlson T, Noren J. Farm tractor fatalities: the failure of voluntary safety standards. *Am J Public Health.* 1979;69(2):146–149.
8. Salmi LR, Weiss HB, Peterson PL, Spengler RF, Sattin RW, Anderson HA. Fatal farm injuries among young children. *Pediatrics.* 1989;83(2):267–271.
9. Simpson SG. Farm machinery injuries. *J Trauma.* 1984;24(2):150–152.
10. Stallones L. Fatal unintentional injuries among Kentucky farm children: 1979 to 1985. *J Rural Health.* 1989;5(3):246–256.
11. Swanson JA, Sachs MI, Dahlgren KA, Tinguely SJ. Accidental farm injuries in children. *Am J Dis Child.* 1987;141:1276–1279.
12. Washington State Dept of Labor and Industries. *The Employers' Guide to Industrial Insurance.* Olympia, Wash: Dept of Labor and Industries; 1986.
13. Washington State Dept of Labor and Industries. *Washington State Work Injury and Illness Summary: 1986.* Olympia, Wash: Dept of Labor and Industries; 1986.
14. US Bureau of Census. *1980 Census of Population: Classified Index of Industries and Occupations.* Washington, DC: US Govt Printing Office; 1980.
15. American National Standards Institute. *American National Standard Method of Recording Basic Facts Relating to the Nature and Occurrence of Work Injuries.* New York, NY: American National Standards Institute; 1985.
16. Schill WJ, McCartin R, Meyer K. Youth employment: its relationship to academic and family variables. *J Vocational Behav.* 1985;26:155–163.
17. State of Washington. *Washington Administrative Code: Chapter 296–131.* Olympia, Wash: State of Washington; 1990.

Public Health Briefs

Impact of HIV Infection on Mortality and Accuracy of AIDS Reporting on Death Certificates

Nancy A. Hessol, MSPH, Susan P. Buchbinder, MD, David Colbert, Susan Scheer, MA, Ronald Underwood, MD, MPH, J. Lowell Barnhart, Paul M. O'Malley, Lynda S. Doll, PhD, and Alan R. Lifson, MD, MPH

ABSTRACT

To assess the impact of HIV infection on mortality and the accuracy of AIDS reporting on death certificates, we analyzed data from 6704 homosexual and bisexual men in the San Francisco City Clinic cohort. Identification of AIDS cases and deaths in the cohort was determined through multiple sources, including the national AIDS surveillance registry and the National Death Index. Through 1990, 1518 deaths had been reported in the cohort and 1292 death certificates obtained. Of the 1292 death certificates, 1162 were for known AIDS cases, but 9% of the AIDS cases did not have HIV infection or AIDS noted on the death certificate. Only 0.7% of the decedents had AIDS listed as a cause of death and had not been reported to AIDS surveillance. AIDS and HIV infection was the leading cause of death in the cohort, with the highest proportionate mortality ratio (85%) and standardized mortality ratio (153 in 1987), and the largest number of years of potential life lost (32 008 years). The devastating impact of HIV infection on mortality is increasing and will require continued efforts to prevent and treat HIV infection. (*Am J Public Health.* 1992;82:561–564)

Introduction

To monitor the impact of specific diseases on mortality in the United States, accurate death certificate information is essential. Such accuracy is important for precise monitoring of mortality trends from acquired immunodeficiency syndrome (AIDS) and possible identification of unrecognized conditions associated with human immunodeficiency virus (HIV) infection. Recently, several US studies, using information from death certificates, have examined the impact of HIV infection and AIDS on mortality.[1–5] These studies are limited by lack of specific risk-group information. Additionally, few studies have been able to assess the accuracy of AIDS reporting on death certificates[5,6] because it is difficult to link US AIDS surveillance and vital statistics registries.

In this analysis, we obtained information on HIV infection, reported AIDS diagnosis, and deaths among a San Francisco cohort of homosexual and bisexual men. Our objectives were to describe underlying causes of death; to assess underreporting of AIDS on death certificates; and to calculate cause-specific proportionate mortality ratios (PMRs), years of potential life lost (YPLL), and standardized mortality ratios (SMRs) for deceased cohort members.

Methods

The San Francisco City Clinic cohort consists of 6704 homosexual and bisexual men recruited between 1978 and 1980 from the municipal sexually transmitted disease clinic for studies of hepatitis B.[7,8] In 1980, hepatitis B vaccine trials began. Unused serum samples from these studies were frozen and stored. In 1983 we began AIDS follow-up studies in this cohort.

All participants reported to have AIDS were considered HIV antibody positive prior to death. For men who died of causes other than AIDS, stored specimens were tested for HIV antibody.

To identify AIDS cases in the cohort, we cross-matched with San Francisco and national AIDS surveillance registries. To identify deaths among cohort members, we gathered information from San Francisco and national AIDS surveillance registries, newspaper obituaries, San Francisco Vital Statistics, and the National Death Index.

Death certificates were obtained from San Francisco Vital Statistics and state registrars. All causes of death and other significant conditions were coded according to the current International Classification of Diseases ninth revision coding manual (ICD-9 CM).[9] According to this system, underlying causes of death are assigned AIDS-related codes only when HIV infection or AIDS is noted. For this analysis, underlying cause of death

Nancy A. Hessol, Susan P. Buchbinder, David Colbert, Susan Scheer, Ronald Underwood, J. Lowell Barnhart, Paul M. O'Malley, and Alan R. Lifson are with the AIDS Office, San Francisco Department of Public Health, San Francisco, Calif. Alan R. Lifson is also with the Department of Epidemiology and Biostatistics, School of Medicine, University of California, San Francisco, Calif. Lynda S. Doll is with the Division of HIV/AIDS, Center for Infectious Diseases, Centers for Disease Control, Atlanta, Ga.

Requests for reprints should be sent to Nancy A. Hessol, MSPH, AIDS Office, 25 Van Ness Ave, Suite 500, San Francisco, CA 94102.

This paper was submitted to the Journal April 25, 1991, and accepted with revisions August 23, 1991.

TABLE 1—Underlying Cause of Death, from Death Certificates, San Francisco City Clinic Cohort, 1978–1990

Cause of death (ICD-9 code)	No.	Percent Known HIV+
Infectious and parasitic diseases (001–139)	1089	99
HIV infection (042–044)	1063	100
Progressive multifocal leukoencephalopathy (046)	3	100
Pneumocystis carinii pneumonia (136)	12	100
Neoplasms (140–239)	40	67
Kaposi's sarcoma (173)	13	100
Lymphoma (202)	14	86
Endocrine, nutritional, and metabolic diseases and immunity disorders (240–279)	24	96
Immune disorders (279)	23	96
Diseases of the blood and and blood-forming organs (280–289)	1	0
Mental disorders (290–319)	2	0
Alcohol dependence syndrome (303)	2	0
Diseases of the nervous system and sense organs (320–389)	3	100
Diseases of the circulatory system (390–459)	9	11
Diseases of the respiratory system (460–519)	10	60
Pneumonia (480–486)	6	67
Diseases of the digestive system (520–579)	19	58
Chronic liver disease and cirrhosis (570–573)	16	50
Congenital anomalies (740–759)	2	0
External causes of injury and poisoning (E800–E999)	91	49
Unintentional injuries (E800–E949)	27	48
Suicide (E950–E959)	43	56
Homicide (E960–E978)	13	46
Injury, undetermined (E980–E989)	8	25
Unknown or unspecified (799.9)	2	0
Total	1292	92

TABLE 2—Underlying Cause of Death on Death Certificates of Known AIDS Cases, without Mention of HIV or AIDS, San Francisco City Clinic Cohort, 1978–1990

Cause of death (ICD-9)	Frequency
Infectious and parasitic diseases (001–139)	23
Progressive multifocal leukoencephalopathy (046)	3
Pneumocystis carinii pneumonia (136)	12
Neoplasms (140–239)	26
Kaposi's sarcoma (173)	13
Lymphoma (202)	11
Endocrine, nutritional, and metabolic diseases and immunity disorders (240–279)	22
Immune disorders (279)	22
Diseases of the nervous system and sense organs (320–389)	3
Diseases of the respiratory system (460–519)	5
Pneumonia (480–486)	4
Diseases of the digestive system (520–579)	6
Chronic liver disease and cirrhosis (570–573)	3
External causes of injury and poisoning (E800–E999)	24
Unintentional injuries (E800–E949)	7
Suicide (E950–E959)	15
Injury, undetermined (E980–E989)	2
Total	109

was determined according to the National Center for Health Statistics guidelines.[10]

Cause-specific PMRs were calculated as the number of deaths from a particular cause divided by the total number of deaths in the cohort. To evaluate temporal trends, cause-specific PMRs were calculated by year of death. The YPLL for each deceased cohort member was calculated by subtracting the age at death from 65 years. Summing the YPLL of all decedents dying of a particular cause gives the cause-specific YPLL. For the years 1981 and 1987, age-adjusted cause-specific death rates within the cohort were compared with United States data,[11] and SMRs were used to measure the ratio of the observed number of deaths to the expected number of deaths.

Results

Through 1990, 1518 deaths had been reported in the cohort and 1292 death certificates obtained. The median year of death was 1987 (range 1978–1990) and the median age at death was 37 years (range 22–76). The racial/ethnic distribution was 92% White, 4% Latino, 3% African American, and 1% other, similar to the distribution of the cohort as a whole.[7,8] Death certificates were obtained from 39 US states and territories. The majority, 1097 death certificates, came from California, including 919 from San Francisco.

The most common underlying cause of death was HIV infection (Table 1). Of the 1292 decedents, 1194 were HIV antibody positive prior to death, 85 were HIV negative at their last blood test, and 13 had no previous HIV antibody results. The percentage of men who were HIV-positive prior to death was highest among those who died of infectious or parasitic diseases and lowest among those who died of liver disease, suicide, or injuries.

Among the 1162 death certificates obtained for men who had previously been reported to AIDS surveillance, 1053 (91%) had HIV or AIDS noted. The most common underlying causes of death in the remaining 9% were neoplasms, immunodeficiency, and infectious and parasitic diseases (Table 2). Among the 1063 decedents with HIV or AIDS noted on the death certificate, only 7 had not been previously reported to AIDS surveillance.

The PMRs for the five most common causes of death are shown in Table 3. The PMR was 84.6% for HIV infection, 3.4% for suicide, and less than 3% for all other causes. Over time, the PMR for HIV infection increased from 0% in 1978–1981 to 14% in 1982, 62% in 1984, 82% in 1986, 88% in 1988, and 93% in 1990.

The cause-specific YPLL was 29 123 for men with HIV or AIDS mentioned on the death certificate, 1251 for suicide, 980 for malignant neoplasms, and less than 900 for all other causes (Table 3). The total YPLL for all AIDS-related deaths, includ-

ing deaths occurring in men identified either through AIDS surveillance or AIDS on the death certificate, was 32 008.

Age-adjusted cause-specific SMRs for the year 1981 show excess mortality for HIV infection and immune disorders (Table 4). SMRs for 1987 reveal a higher than expected rate for HIV infection and immune disorders and for suicide.

TABLE 3—Proportionate Mortality Ratios (PMRs) and Years of Potential Life Lost (YPLL) for the Five Leading Causes of Death, San Francisco City Clinic Cohort, 1978–1990

Cause of death (ICD-9 code)	PMR	YPLL
HIV infection (042–044)	84.6	29 123
Suicide (E950–E959)[a]	3.4	1 251
Malignant neoplasms (140–208)[a]	2.9	980
Unintentional injuries (E800–E949)[a]	2.1	830
Other infectious and parasitic diseases (001–041, 045–139)[a]	2.1	729
AIDS-related subtotal (includes all AIDS cases)		32 008
Total, all causes		35 546

[a]Includes reported AIDS cases

TABLE 4—Age-Adjusted Cause-Specific Standardized Mortality Ratios (SMRs) and 95% Confidence Intervals (CIs) for Years 1981 and 1987, San Francisco City Clinic Cohort

Cause of Death	SMR (95% CI) 1981	SMR (95% CI) 1987
HIV infection and immune disorders[a]	127.7 (3.2–696)	153.5 (131–179)
Suicide	0.6 (.02–3.5)	3.4 (1.1–7.9)
Malignant neoplasms	1.3 (.15–4.6)	1.1 (.2–3.1)
Unintentional injuries	0.4 (.05–1.6)	0.3 (.01–1.7)
Other infectious and parasitic diseases	6.7 (.17–37.1)	0

95% confidence intervals were calculated using the Poisson distribution.[14]
[a]Due to temporal changes in both the ICD-9 coding system and the Centers for Disease Control AIDS case definition, these two leading causes of death were combined.

Discussion

In the 1986–1987 US population, injuries and heart disease were the most common underlying causes of death among white men aged 25 to 44.[11] Our findings from the San Francisco City Clinic cohort show a different pattern, with AIDS as the leading underlying cause of death and YPLL. The underlying causes of death in our cohort are, however, similar to those of the 1986–1987 San Francisco population for white men aged 25 to 44.[11]

Unlike mortality studies of HIV infection in intravenous drug users,[4] most underlying causes of death associated with HIV infection in this cohort met the AIDS case definition. For example, sepsis was reported to be a common cause of death in intravenous drug users who did not have AIDS,[4] but in our study sepsis was usually reported with AIDS. Excess rates of suicide, however, warrant further investigation.

AIDS surveillance provided more complete AIDS reporting than did death certificates, with 0.7% versus 9% underreporting. This low rate of underreporting (0.7%) is likely due to comprehensive case surveillance in San Francisco, which includes death certificate review for case identification.

Even in a San Francisco cohort of homosexual and bisexual men, AIDS is underreported as an underlying cause of death on death certificates. The omission of HIV infection or AIDS on the death certificates may be due to lack of knowledge about proper completion of death certificates, intentional cover-ups to protect patients' confidentiality, cause of death's being unrelated to HIV infection, or incomplete knowledge of the decedent's medical condition.

The impact of HIV infection and AIDS on premature mortality will continue to increase both in the cohort and in the United States. Because homosexual and bisexual men were among the first risk groups infected with HIV,[7] the patterns seen in the San Francisco City Clinic cohort may predict future patterns for the United States. National mortality trends have already changed, with HIV infection ranked sixth in 1989 for estimated YPLL among all persons dying before age 65.[12] In 1988, HIV infection/AIDS became the third leading cause of death among US men 25 to 44 years of age; it was estimated to be the second leading cause of death in 1989.[13] As HIV infection spreads into other transmission risk groups, such as intravenous drug users, their sex partners, and their children, AIDS and HIV infection will continue to be an important cause of morbidity and premature mortality in the United States. □

Acknowledgments

This study was supported by cooperative agreement number U64/CCU900523-06 from the Centers for Disease Control, Atlanta, Ga.

Preliminary findings of this study were presented at the 118th meeting of the American Public Health Association in New York City on October 2, 1990.

We thank the following people for their assistance in this study: Torsten Bodecker, Lois Conley, Lyn Fischer-Ponce, Aida Flandez, Delia Garcia, Sandra Hernandez, Scott Holmberg, Fred Ingram, Mitchell Katz, Robert Kono, Irene Lee, Bobby Martin, Marvin Montenegro, Frank Phillips, Michael Piccini, Martha Rogers, Chris Rubino, George Rutherford, Belinda Van, Robert Wade, Katie Young, and the men in the San Francisco City Clinic cohort study.

References

1. Buehler JW, Devine OJ, Berkelman RL, Chevarley FM. Impact of the human immunodeficiency virus epidemic on mortality trends in young men, United States. *Am J Public Health.* 1990;80:1080–1086.
2. Saunders LD, Rutherford GW, Lemp GF, Barnhart JL. Impact of AIDS on mortality in San Francisco, 1979–1986. *J AID Syndr.* 1990;3:921–924.
3. Menendez BS, Drucker E, Vermund SH, et al. AIDS mortality among Puerto Ricans and other Hispanics in New York City, 1981–1987. *J AID Syndr.* 1990;3:644–648.
4. Selwyn PA, Hartel D, Wasserman W, Drucker E. Impact of the AIDS epidemic on morbidity and mortality among intravenous drug users in a New York City methadone maintenance program. *Am J Public Health.* 1989;79:1358–1362.
5. Kristal AR. The impact of the acquired immunodeficiency syndrome on patterns of premature death in New York City. *JAMA.* 1986;255:2306–2310.
6. McCormick A. Trends in mortality statistics in England and Wales with particular reference to AIDS from 1984 to April 1987. *Br Med J.* 1988;296:1289–1292.
7. Hessol NA, Lifson AR, O'Malley PM, Doll LS, Jaffe HW, Rutherford GW. Prevalence, incidence, and progression of human immunodeficiency virus infection in homosexual and bisexual men in hepatitis B vaccine trials, 1978–1988. *Am J Epidemiol.* 1989;130:1167–1175.

8. Rutherford GW, Lifson AR, Hessol NA, et al. Course of HIV-1 infection in a cohort of homosexual and bisexual men: an 11-year follow-up study. *Br Med J.* 1990;301:1183–1188.
9. *International Classification of Diseases.* 9th revision. Washington, DC: US Government Printing Office; 1989. Clinical modification, DHHS publication PHS-89-1260.
10. National Center for Health Statistics. *Instruction Manual Part 2a: Instructions for classifying the underlying cause of death, 1990.* Hyattsville, Md: US Dept of Health and Human Services, Public Health Service; 1990.
11. National Center for Health Statistics. *Compressed Mortality, 1979–1987.* National Vital Statistics System. Hyattsville, Md: US Dept of Health and Human Services, Public Health Service, Centers for Disease Control; 1990.
12. Centers for Disease Control. Update: years of potential life lost before age 65—United States, 1988 and 1989. *MMWR.* 1991;40:60–63.
13. Centers for Disease Control. Mortality attributable to HIV infection/AIDS—United States, 1981–1990. *MMWR.* 1991;40:41–44.
14. Gardner MJ, Altman DG. *Statistics with Confidence.* London: British Medical Journal; 1989:59.

Seroprevalence of HIV-Type 1 in a Northern California Health Plan Population: An Unlinked Survey

Robert A. Hiatt, MD, PhD, Frank J. Capell, and Michael S. Ascher, MD

ABSTRACT

We conducted an unlinked seroprevalence survey for human immunodeficiency virus–type 1 (HIV-1) of 9821 persons who had a routine personal health appraisal examination in 1989 while members of the Kaiser Permanente Medical Care Program in northern California. An outside laboratory performed enzyme immunoassay (EIA) analyses, and the California Viral and Rickettsial Diseases Laboratory confirmed EIA-reactive samples by immunofluorescent assay and Western blot assay. Only 20 specimens (0.2%) were confirmed as positive, and 18 were from men. These data suggest that, at the time of this survey, HIV-1 infection was not widespread in the northern California population represented by this health plan membership. (*Am J Public Health.* 1992;82:564–567)

Introduction

Seroprevalence studies designed to assess the spread of human immunodeficiency virus–type 1 (HIV-1) infection in the general population are usually biased by the selected or voluntary nature of the sample or are limited in their generalizability because of the specialized nature of the survey population.[1–5] Unlinked (blinded) surveys bypass the need for informed consent and thus eliminate selection bias from this source.

We conducted an unlinked survey among nonhospitalized members of a large northern California prepaid health plan with characteristics representative of the general population in that region. The study protocol was approved by the Institutional Review Board of the Kaiser Permanente Medical Care Program (KPMCP), Northern California Region. We believe our results can be generalized to the symptom-free general population surrounding the city of San Francisco, one of the areas with the highest prevalence of HIV-1 in the United States.

Methods

Study Population

The KPMCP is a group-model, prepaid health maintenance organization that in 1989 provided comprehensive medical care to 2 190 629 members, or approximately 30% of the population in the counties of northern California where KPMCP provides services. The membership is comparable with the general population in terms of age and sex, and it includes the full spectrum of the region's racial, occupational, and socioeconomic groups.[6,7]

Data Collection

At most KPMCP facilities, members older than 15 years of age can voluntarily take a personal health appraisal (PHA) that includes a questionnaire, a physical examination, and laboratory tests. A PHA might be done because of a member's or provider's perceived need for a checkup, either to collect baseline data when the member first joins the health plan or starts with the new provider, or to provide a broadly based assessment for nonspecific complaints or concerns. It is generally not used as a diagnostic tool for focused complaints. Persons who take PHAs are likely to differ from the general plan membership; however, except in one facility,[7] this trend has not been systematically examined.

Aliquots of sera from 10 000 consecutive PHA samples drawn at nine Kaiser Foundation facilities (Figure 1) in October

Robert A. Hiatt is with the Division of Research, Kaiser Permanente Medical Care Program. Frank J. Capell is with the Office of AIDS and Michael S. Ascher is with the Viral and Rickettsial Disease Laboratory of the California Department of Health Services.

Requests for reprints should be sent to Robert A. Hiatt, MD, PhD, Division of Research, Kaiser Permanente Medical Care Program, 3451 Piedmont Avenue, Oakland, CA 94611–5463.

This paper was submitted to the Journal April 17, 1991, and accepted with revisions September 20, 1991.

and November of 1989 were coded and linked with KPMCP data on the member's 5-year age group, sex, zip code, date of PHA, and facility. The coded aliquots were sent on dry ice to an outside non-KPMCP laboratory (Alpha Therapeutics in Memphis, Tenn). All specimens were tested by enzyme immunoassay (EIA) (Pharmacia, ENI-Virgo), for which this laboratory is federally licensed, and positive specimens were retested. Specimens found by EIA to be positive for HIV-type 1 and an equal number of randomly selected negative specimens were sent to the Viral and Rickettsial Disease Laboratory of the California Department of Health Services in Berkeley, where repeat EIAs, Western blot assays, and immunofluorescent antibody testing were done. Codes for specimens testing positive by EIA, Western blot, and immunofluorescent antibody assay were reported as HIV-1 positive to the Department of Health Services' Office of AIDS, which also received coded demographic information from the KPMCP. It was not possible for any party in the study to trace a positive result to any individual.

Results

From the original dataset of 10 000 samples, 179 samples were eliminated because they were either found to have been duplications or drawn for specific laboratory tests rather than for PHAs. Age and sex distribution for the remaining 9821 samples was similar to that of all health plan members older than 15 years of age (Table 1). Twenty of the 9821 samples were confirmed as positive for HIV-type 1, for an overall prevalence of 0.2%. Eighteen of these were from men, and the age-specific prevalence was highest among those 20 to 29 years of age (Table 2). The distribution by facility or zip code grouped by major geographic area in northern California suggests that the prevalence of HIV-1 was slightly higher for persons living on the San Francisco peninsula and in the East Bay than for persons living in other areas served by the health plan, but numbers were small and 95% confidence intervals were wide.[8]

The San Francisco peninsula sample included 419 residents of San Francisco who obtained PHAs from the South San Francisco facility, but the sample did not include specimens from the San Francisco facility, which does not have a PHA program. Multiple logistic regression analyses for age, sex, and regions separately and together confirmed that the strongest determinant of HIV-1 positivity in this sample was being male and 20 to 29 years of age. No marked differences by region were found.

FIGURE 1—Location of principal Kaiser Foundation hospitals contributing serum specimens to survey.

TABLE 1—Distribution of Survey Sample Compared with Kaiser Permanente Medical Care Program (KPMCP) Membership in December 1989, by Age and Sex

	Survey		KPMCP	
Age, y	No.	(%)	No.	(%)
15–19	183	(1.9)	152 276	(8.9)
20–29	1181	(12.0)	309 798	(18.2)
30–39	2176	(22.2)	386 504	(22.7)
40–49	2217	(22.6)	339 223	(19.9)
50–59	1690	(17.2)	219 360	(12.9)
60–69	1489	(15.1)	179 919	(10.5)
70+	885	(9.0)	117 124	(6.9)
Total	9821	(100.0)	1 704 204	(100.0)
% Female		56.9		52.0

Discussion

Our study illustrates the usefulness of unlinked seroprevalence surveys in providing important public health information while protecting the confidentiality of the individual.[9–11] Such surveys are methodologically required in states such as California, where the results of individual HIV-1 testing are kept confidential by law, but they also have advantages in states without such laws, where any survey based on voluntary participation is likely to result in selection bias.[12] Considering that persons at high risk of HIV-1 infection

TABLE 2—HIV-1 Seropositivity per 1000 Members by Age, Sex, and Region

Characteristic	No. Tested	No. Positive	Rate per 1000	95% Confidence Interval[a]
Age, y				
< 20	183	0	0.0	—
20–29	1181	6	5.1	1.9–11.1
30–39	2176	6	2.8	1.0–6.0
40–49	2217	4	1.8	0.5–4.6
50–59	1690	2	1.2	0.1–4.3
60–69	1489	1	0.7	0.02–3.7
70+	885	1	1.1	0.03–6.3
Sex				
Male	4236	18	4.2	2.5–6.7
Female	5585	2	0.4	0.04–1.3
Residence				
North Bay	1935	4	2.1	0.6–5.3
East Bay	4228	10	2.4	1.1–4.4
San Francisco peninsula	862	2	2.3	0.3–8.4
South Bay	2500	4	1.6	0.4–4.1
Other	296	0	0.0	—
Total	9821	20	2.0	1.2–3.2

[a]For the expectation of a Poisson variable.

were not excluded from our study, a prevalence of 0.2% suggests that, by the end of 1989, HIV-1 had not spread to any substantial degree to members of this health plan population who were not at high risk.

Persons from San Francisco, who may have been at higher risk than those from other areas served by the health plan, were underrepresented in this survey because the San Francisco medical center did not have a PHA program. The survey should thus be viewed as reflecting levels of HIV-1 infection in a primarily urban area surrounding, but not including, San Francisco. Moreover, intravenous drug users are not likely to have been included in this survey. With these caveats in mind, we believe this sample still more closely represents the overall level of HIV-1 infection in the general population of the San Francisco Bay Area than have previous surveys.

Estimates of HIV-1 seroprevalence among heterosexuals have been derived mainly from surveys of high-risk or self-selected volunteers, such as blood donors,[13] military recruits,[1,14] active duty personnel,[2,15] prisoners,[3] outpatients or emergency patients,[16] and those from sexually transmitted disease clinics.[5] A blinded serosurvey of all 135 808 childbearing women in California during the third quarter of 1988 found 0.074% positive for HIV-1; 0.129% of San Francisco Bay Area childbearing women were positive[4] compared with 0.036% of the health plan women in our survey.

The closest comparison data to our survey are the results of the Centers for Disease Control's anonymous seroprevalence surveys at sentinel hospitals in the United States.[10] For eight hospitals in the West surveyed from January 1988 through June 1989, the overall prevalence was 0.35% (range: 0.1–0.7%). The peak prevalence was highest in the 25- to 34-year-old age group. Persons admitted for diagnoses often associated with HIV-1 infection were excluded.[17] Nevertheless, the sentinel hospitals' slightly higher overall prevalence of 0.35%, compared with a prevalence of 0.20% found by our survey, could be due to the higher probability of hospitalized patients being infected with HIV.

The results of our survey may be reassuring to persons concerned about the heterosexual spread of HIV-1 infection. At least in this northern California health plan population, levels of infection appear low. However, this does not minimize the need for educational efforts to optimize the practice of safe sexual behavior. The most conservative projections indicate that the number of persons with active HIV-1 infection will be substantial throughout at least the next decade,[18] and even a low level of risk for this uniformly fatal disease is important. □

Acknowledgments

The research was supported in part by the California State Department of Health Services, Office of AIDS Contract No. 88-95042.

This paper was presented at the 118th Annual Meeting of the American Public Health Association, September 31–October 4, 1990, New York City.

References

1. Burke DS, Brundage JF, Herbold JR, et al. Human immunodeficiency virus infections among civilian applicants for United States military service, October 1985 to March 1986. *N Engl J Med.* 1987;317:131–136.
2. Kelley PW, Miller RN, Pomerantz R, Wann F, Brundage JF, Burke DS. Human immunodeficiency virus seropositivity among members of the active duty US Army 1985–1989. *Am J Public Health.* 1990;80:405–410.
3. Singleton JA, Perkins CI, Trachtenberg AI, Hughes MJ, Kizer KW, Ascher M. HIV antibody seroprevalence among prisoners entering the California correctional system. *West J Med.* 1990;153:394–399.
4. Capell FJ, Vugia DJ, Mordaunt V, et al. Distribution of human immunodeficiency virus type 1 infection in childbearing women in California. *Am J Public Health.* 1992;82:254–256.
5. Del Tempelis C, Shell G, Hoffman M, Benjamin RA, Chandler A, Francis DP. Human immunodeficiency virus infection in women in the San Francisco Bay Area. *JAMA.* 1987;258:474–475. Letter.
6. Hiatt RA, Friedman GD. The frequency of kidney and urinary tract diseases in a defined population. *Kidney Int.* 1982;22:63–68.
7. Krieger N. Overcoming the absence of social class data in medical records: a census-based methodology to assess the demographic features of the membership of a large pre-paid health plan. Proceedings of the 118th annual meeting of the American Public Health Association; September 30–October 4, 1990; New York, NY.
8. Box GEP, Hunter WG, Hunter JS. *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building.* New York, NY: John Wiley; 1978:644.
9. Pappaioanou M, Dondero TJ Jr, Petersen LR, Onorato IM, Sanchez CD, Curran JW. The family of HIV seroprevalence surveys: objectives, methods, and uses of sentinel surveillance for HIV in the United States. *Public Health Rep.* 1990;105:113–119.
10. St Louis ME, Olivo N, Critchley S, et al. Methods of surveillance for HIV infection at US sentinel hospitals. *Public Health Rep.* 1990;105:140–146.
11. Petersen LR, Calonge NB, Chamberland ME, Engel RH, Herring NC. Methods of surveillance for HIV infection in primary care outpatients in the United States. *Public Health Rep.* 1990;105:158–162.
12. Hull HF, Bettinger CJ, Gallaher MM, Keller NM, Wilson J, Mertz GJ. Comparison of HIV-antibody prevalence in patients consenting to and declining HIV-antibody testing in an STD clinic. *JAMA.* 1988;260:935–938.
13. Hughes MJ, Winter SL, Perkins CL, Kizer KW, Capell FJ, Trachtenberg AL. Prevalence of HIV antibody among blood donors in California. *N Engl J Med.* 1989;321:974–975.
14. Trends in human immunodeficiency virus infection among civilian applicants for mil-

itary service: United States, October 1985–March 1988. *MMWR.* 1988;37:677–679.
15. Prevalence of human immunodeficiency virus antibody in US active-duty military personnel, April 1988. *MMWR.* 1988;37:461–463.
16. Baker JL, Kelen GD, Sivertson KT, Quinn TC. Unsuspected human immunodeficiency virus in critically ill emergency patients. *JAMA.* 1987;257:2609–2611.
17. St Louis ME, Rauch KJ, Petersen LR, Anderson JE, Schable CA, Dondero TJ. Seroprevalence rates of human immunodeficiency virus infection at sentinel hospitals in the United States. *N Engl J Med.* 1990;323:213–218.
18. Estimates of HIV prevalence and projected AIDS cases: summary of a workshop, October 31–November 1, 1989. *MMWR.* 1990;39:110–119.

The Validity of Self-reported HIV Antibody Test Results

Jane McCusker, MD, DrPH, Anne M. Stoddard, ScD, and Ellen McCarthy, MPH

ABSTRACT

Three hundred twenty-six participants in a multisite study were initially tested for human immunodeficiency virus (HIV) antibody and later interviewed; 228 were recent drug injectors and 38 were HIV positive. Later, 264 (81%) correctly reported their test results, 9 (3%) gave incorrect results, and 53 (16%) said either that they had not been tested or that they did not know the results. The predictive values of positive and negative self-reports were 90% and 98%, respectively ($P < .01$). (*Am J Public Health.* 1992;82:567–569)

Introduction

Research on the occurrence and correlates of human immunodeficiency virus type 1 (HIV-1, referred to in this paper as HIV) ideally is based on laboratory-confirmed antibody test results. In some situations, however, laboratory confirmation is not available—perhaps because of concern with confidentiality or inadequate resources to test all study subjects—and only self-reported results are available. The decision on whether to use these results rests on whether their inclusion will introduce bias. The inclusion of subjects who report inaccurate results would introduce information bias; their exclusion might introduce selection bias, for example, by excluding subjects who are not retested because they report a positive test result. In this study we investigated the validity of self-reported results among individuals who were initially tested in a multisite HIV surveillance project targeting intravenous drug users and their sexual partners,[1,2] and who were later reinterviewed regarding their test results.

Methods

Methods of enrollment and data collection have been reported previously.[2,3] Study participants included persons requesting HIV antibody testing at two city health department clinics, two community health centers, and a men's jail from September 1987 to December 1989, or at three drug abuse treatment programs during 1988 and 1989. Interview data were kept in client records at drug treatment sites. At other sites, interview data were used only by the HIV counselor and were stored confidentially. Participants were informed about these procedures. All testing was carried out in one laboratory: commercially available FDA-approved enzyme-linked immunosorbent assays (ELISA) kits, were used for initial serological screening and positives were confirmed with a Western blot test.

For this analysis, we examined the subset of participants who were reinterviewed during the study period. Reasons for reinterview included retesting at the same or another site and subsequent admission (or readmission) to one of the drug treatment programs. Self-reported results at the reinterview were compared with actual test results by HIV antibody status and other characteristics of the subjects. Differences in percentages were tested by the normal approximation to the binomial distribution, using a 5% level of statistical significance. Predictive values of positive and negative self-reports were computed.[4]

Results

A total of 2304 individuals were tested for HIV antibody, of whom 33% were recent needle users and 22% were tested while in drug abuse treatment. Of the 2304 tested, 326 (14%) were reinterviewed at the same or another study site

Jane McCusker and Anne M. Stoddard are with the University of Massachusetts AIDS Research Unit, University of Massachusetts School of Public Health, Amherst, Mass. Ellen McCarthy is with Spectrum Inc, Westborough, Mass.

Requests for reprints should be sent to Jane McCusker, MD, DrPH, School of Public Health, Arnold 406, University of Massachusetts, Amherst, MA 01003.

This paper was submitted to the Journal April 16, 1991, and accepted with revisions August 26, 1991.

TABLE 1—Agreement between Actual and Self-reported HIV-1 Antibody Status, Stratified by Selected Variables

Variable	Total	Result Reported			Correct Result[b]		
		No.	%	P[a]	No.	%	P[a]
Actual HIV result							
Positive	38	33	87	.58	27	82	<.001
Negative	288	240	83		237	99	
Recent IVDU							
Yes	228	190	83	.82	182	96	.18
No	96	81	84		80	99	
Interval from test							
<6 Months	167	136	81	.25	132	97	.63
6+ Months	159	137	86		132	96	
Follow-up site							
Same	175	150	86	.29	149	99	.004
Different	150	122	81		114	93	
Treatment	206	167	81	.09	158	95	.02
Other	119	105	88		105	100	

[a]*P* values are based on normal approximation to the binomial distribution.
[b]The denominator is the number of those reporting a result.

during the study period. Those reinterviewed were significantly more likely to be recent needle users (70%) and to have been tested initially during drug abuse treatment (41%).

The HIV seroprevalence in the study group was 12% (38/326). At the second interview 21 subjects (6%) said they had not been tested before, 32 (10%) said they had been tested but did not know the result, and 273 (84%) reported at least one prior test result. In the latter group we tried to match the report to the index test by searching for the closest self-reported test date. Of the 221 subjects who gave a month and year, we found a report for the same month and year for 107, a report within 1 month for 54, and a report dated more than 1 month from that given for 60. Of the 39 who gave only a year, we found a report in the same year for 24 and a report in a different year for 15. The remaining 13 gave no test date.

The percentages of HIV-seropositive and -seronegative subjects who reported test results were similar (87% and 83%, respectively), but among those who reported results, seropositive subjects were significantly less likely than seronegative subjects to report their results correctly. Of 243 negative test reports, 6 were false negatives, whereas 3 of 30 positive reports were false positives, giving predictive values of 98% for a negative report and 90% for a positive report. We are confident that subjects were reporting the same result as the index test in all of the false-positive reports and all but one of the false-negative reports. (See Appendix A for details of these cases.)

The correspondence of reported with actual test results was stratified by several other variables (Table 1). The percentage of missing and unknown results did not differ significantly by recent intravenous drug use, interval between the two visits, whether the follow-up was at a drug abuse treatment site or another site, or whether the follow-up was at the same site as the initial visit. There was no significant association between accuracy of the reported result and recent intravenous drug use or interval between visits. Test results were more likely to be reported correctly if individuals were reinterviewed at the site where they had been tested and if this site was a non–treatment site. The sample was too small for us to determine whether these site effects were independent. Furthermore, for most of the subjects interviewed at a different site, this was a treatment site (88%).

Discussion

We found that HIV-seropositive persons were more likely to misreport their test results than were HIV-seronegative persons, particularly if they had been tested at a different site. We were not able to determine the reasons for this misreporting or whether it was intentional. During the study period, drug users who tested positive were given priority for admission into treatment, which might have accounted for the false-positive results. On the other hand, false-negative results might have been reported by individuals who wished their serostatus to be confidential for fear of stigmatization.

There was a strong trend for those who were followed up at the same site where they were tested to report their results more correctly, although this trend was of borderline statistical significance. This finding supports previous findings that intravenous drug users are more likely to provide correct information when they know that other sources of information exist.[5]

We acknowledge the limited generalizability of the results of this study, which was restricted to individuals who initially chose to be tested and were recontacted either because they wished to be retested or because they entered drug abuse treatment. However, these limitations apply to most research on self-reported HIV status among drug-using populations.

Previous research on drug users has generally found good agreement between self-reports and official records,[6,7] although this finding is not universal.[8] It is important to distinguish between inability to give correct information, owing to memory lapses or recent drug use, from unwillingness; social desirability and the perceived acceptability of the response can be powerful predictors of validity.[5]

The predictive value of test results is of interest to those deciding whether to include self-reported results in estimates of total seroprevalence. Our data suggest that the predictive values of both positive and negative self-reports were high, with positive self-reports somewhat less predictive than negative self-reports (90% vs 98%). Predictive values in other populations will vary with HIV seroprevalence.[4]

Given their generally high validity, it may be appropriate to use self-reported results in certain situations, for example, when a substantial proportion of the population does not have laboratory-confirmed results and where the costs of potentially redundant testing are prohibitive. In these situations, the selection bias introduced by exclusion of self-reported results may be greater than the information bias introduced by their inclusion. To minimize information bias, it would be important to assure clients in treatment programs that reporting would be confidential. Seroprevalence estimates that are based on self-reported test results may also need to be adjusted for

seroconversion, particularly if there has been a substantial lag since the most recent negative test. □

Acknowledgments

This work was supported by grants from the Commonwealth of Massachusetts, Department of Public Health, and the National Institute on Drug Abuse (R01-DA05615-01A2).

The authors wish to thank the members of the Worcester AJDS Consortium for their support in this effort, especially the interviewers and the field coordinators for collecting the data used in this study; and Joan Landon, MPH, who assisted with the analysis.

References

1. Centers for Disease Control. Coordinated community programs for HIV prevention among intravenous drug users—California, Massachusetts. *MMWR.* 1989;38:369–374.
2. McCusker J, Koblin B, Lewis BF, Sullivan J. Demographic characteristics, risk behavior, and HIV seroprevalence among intravenous drug users by site of contact: results from a community-wide HIV surveillance project. *Am J Public Health.* 1990;80:1062–1067.
3. Koblin BA, McCusker J, Lewis BF, Sullivan JL. Racial/ethnic differences in HIV-1 seroprevalence and risky behaviors among intravenous drug users in a multi-site study. *Am J Epidemiol.* 1990;132:837–846.
4. Mausner JS, Kramer S. *Epidemiology An Introductory Text.* 2nd ed. Philadelphia, Pa: WB Saunders Co; 1985:chap 9.
5. Harrell AV. Validation of self-report: the research record. In: *Self-report Measures of Estimating Drug Use.* Rockville, Md: US Dept of Health and Human Services; 1985. NIDA Research Monograph 57: 12–21.
6. Stephens R. The truthfulness of addict respondents in research projects. *Int J Addict.* 1972;7:549–558.
7. Bonito AJ, Nurco DN, Shaffer JW. The veridicality of addicts' self-reports in social research. *Int J Addict.* 1976;11:719–724.
8. Smart RG, Jarvis GK. Do self-report studies of drug use really give dependable results? *Can J Criminology.* 1981;23:3–12.

APPENDIX A—Further Details on the False-Positive and False-Negative Reports

False positives (n = 3). Two of these subjects reported having been tested in the same month and year as the index test. Both subjects were retested at the reporting visit, with another negative test result. The third subject reported only a year of testing (the same as the index test year) but reported three consecutive positive test results during that year. This subject was not retested and no validated reports of prior positive test results were obtained.

False negatives (n = 6). Four of these subjects reported a test date that was either the same as the index test date or more recent. The fifth subject reported two negative test results during the same year as the index test date, but gave no month. However, at a later interview, this person reported negative test results for the same year as well as for the following year. Inasmuch as documented cases of seroreversion from HIV seropositive to HIV seronegative are extremely rare, these cases in all probability reflect inaccurate reporting. The sixth subject reported a negative test result that was 2 years before the index positive result. We were not able to determine whether this reflected inaccurate reporting of the test result, or failure to report on the index test.

Cognitive and Attitudinal Impacts of a University AIDS Course: Interdisciplinary Education as a Public Health Intervention

Ronald P. Strauss, DMD, PhD, Inge B. Corless, RN, PhD, James W. Luckey, PhD, Charles M. van der Horst, MD, and Betty H. Dennis, PharmD, MS

ABSTRACT

This paper describes an interdisciplinary, variable credit-bearing university course on acquired immunodeficiency syndrome (AIDS) that enrolled 429 students. Pre- and post-course questionnaires were used to assess knowledge and attitudes relative to AIDS and these were compared to National Health Interview Survey findings. Considerable cognitive and attitudinal changes occurred over the course period. University courses, taught annually, were found to be an efficient mechanism for educating large numbers of future community leaders and professionals about AIDS. (*Am J Public Health.* 1992;82:569–572)

Introduction

The onset and spread of human immunodeficiency virus (HIV) infection has raised many social, ethical, and clinical issues that are of importance to individuals from varied backgrounds. A university course that provides a detailed education about HIV infection and acquired immunodeficiency syndrome (AIDS) serves to develop expertise in future community and professional leaders, and may enlighten society's response to this health care and social crisis. The educational topics incorporated in a disease-specific educational offering are applicable beyond the focus on AIDS. AIDS was dealt with in this course not only as a disease, but also as a case study of how societies deal with contagion, stigma, disability, death, social stratification, and access to scarce resources.

Ronald P. Strauss is with the School of Dentistry, Inge B. Corless is with the School of Nursing, James W. Luckey is with the School of Public Health, Charles M. van der Horst is with the School of Medicine, and Betty H. Dennis is with the School of Pharmacy at the University of North Carolina at Chapel Hill.

Requests for reprints should be sent to Dr. Ronald P. Strauss, University of North Carolina School of Dentistry, CB #7450, Chapel Hill, NC 27599.

This paper was submitted to the Journal April 2, 1991, and accepted with revisions November 27, 1991.

TABLE 1—Risk Factor Knowledge: National, Precourse, and Postcourse Findings

	Percent Responding "Very or Somewhat Likely"		
	NHIS Ages 18–29[a] (n = 2194)	Precourse (n = 339)	Postcourse (n = 245)
Question: How likely do you think it is that a person will get AIDS or the AIDS virus infection from:			
Living near a hospital or home for AIDS patients	5	0	0
Working near someone with the AIDS virus	11	1	0
Eating in a restaurant where the cook has the AIDS virus	13	6	4
Kissing—with exchange of saliva—a person who has the AIDS virus	51	26	15
Shaking hands with, touching, or kissing on the cheek someone who has the AIDS virus	8	0	0
Sharing plates, forks, or glasses with someone who has the AIDS virus	26	7	2
Using public toilets	17	4	1
Sharing needles for drug use with someone who has the AIDS virus	98	99	99
Being coughed or sneezed on by someone who has the AIDS virus	23	8	2
Attending school with a child who has the AIDS virus	6	0	0
Mosquitoes or other insects	30	10	4

[a]Health Information Survey data from 1989 provisional estimates, January to March 1989.

TABLE 2—Individual Risk Factors and HIV Testing Experience: National, Precourse, and Postcourse Findings

	Those Responding "Yes"		
	NHIS Ages 18–29[a] (n = 2194)	Pre-course (n = 339)	Post-course (n = 245)
	%	%	%
Did you answer "yes" to having any one of 6 possible HIV risk statements (NHIS)?	3	6.4	9.5
Have you ever had your blood tested for the AIDS virus?	26	18	23
If yes, was the test			
part of a blood donation	67	60	60
part of a blood transfusion	0	1	4
voluntarily sought	19	24	32
part of some other activity that requires a blood sample	16	21	19

[a]Health Information Survey data from 1989 provisional estimates, January to March 1989.

Universities in the United States have responded to AIDS by incorporating relevant information into courses currently offered to students and by developing new courses.[1–4] Cross-sectional surveys and other reports have defined gaps in HIV-related knowledge and attitudes among college students.[5–27] However, the literature on the impact of AIDS education suggests that knowledge and attitude changes do occur over the course of educational programs.[1–3,28–31]

This paper describes a study of an AIDS course as a public health intervention and examines cognitive and attitudinal changes that occurred in students over the course period.

The course, "AIDS: Principles, Practices and Politics," combined faculty and resources from health and social science disciplines. One faculty member from the Schools of Dentistry, Medicine, Nursing, Pharmacy, and Public Health participated in organizing and teaching the course, thus exemplifying the interdisciplinary approach necessary to provide care and resolve the AIDS crisis.

Students, who were drawn from every school on the university campus, included undergraduates (69%), professional/graduate students (28%), and others (3%). Offered at one to three credit levels, the course involved weekly lectures, readings, and a graduate seminar. Enrollment between 1989 and 1992 has varied from 250 to 550 students per semester, with 429 enrolled in 1989.

Methods

In 1989, an anonymous precourse questionnaire was administered to students to examine their risk factors, knowledge, and attitudes. Three hundred and thirty-nine students took the pretest, and their responses were compared with those of the 245 students who took the posttest. The data were treated in the aggregate, as the anonymity of the instrument precluded matching individual precourse and postcourse questionnaires. Some of the questions were drawn from the National Health Interview Survey (NHIS),[32] and findings include comparisons with these national data. Postcourse test mean responses were compared with precourse means using *t* tests and chi-square.

Results

Student knowledge about the risk factors for infection with the AIDS virus are displayed in Table 1. When compared with national surveys of US persons aged 18 to 29, from the same time period (NHIS, January to March 1989),[32] it is evident that our sample of university students was less concerned about casual HIV transmission.

Students starting this course were more educated and accurate about risk factors for HIV infection than those in the national sample. Though students expressed no concern about living near, shaking hands with, touching, kissing on the cheek, or attending school with persons infected with the AIDS virus, the precourse survey showed some concern about other forms of casual contagion. These included sharing eating utensils (7%), eating in a restaurant where the cook has the virus (6%), and being coughed or sneezed on by an infected person (8%). There was also some concern about mosquitoes and other insects (10%). The relatively low level of precourse concern or misinformation was reduced during the course.

TABLE 3—Precourse and Postcourse Attitudes about Specific Issues

	Those Responding "Positive"					
	Precourse (n = 339)		Postcourse (n = 245)		Chi-Square (*df* = 1)	*P*
	%	#	%	#		
Health Care						
Should doctors be allowed to refuse to care for someone who has the AIDS virus?	27.8	(93)	14.1	(34)	15.202	.000***
Should nurses be allowed to refuse to care for someone who has the AIDS virus?	29.3	(98)	15.8	(38)	14.134	.000***
Should health care workers be notified if a patient has tested positive for the AIDS virus?	90.2	(303)	84.4	(206)	4.354	.037*
Should health officials be allowed to distribute information that contains sexually explicit language as part of AIDS prevention efforts?	91.3	(306)	96.3	(234)	5.625	.018*
Should health officials be allowed to distribute information that contains sexually explicit drawings as part of AIDS prevention efforts?	81.2	(272)	90.5	(220)	9.703	.002*
Should public health and medical workers take the responsibility to notify the spouse or sexual partner(s) of those who test positive for the AIDS virus?	76.3	(254)	81.5	(194)	2.252	.133
Schools						
Should children who have the AIDS virus be allowed to go to school with other children?	94.9	(316)	96.7	(235)	1.112	.292
If you had a school-age child, would you worry about him or her catching AIDS from a classmate?	34.9	(116)	25.5	(62)	5.832	.016*
Employment						
Should coworkers be allowed to refuse to work near someone who has the AIDS virus?	38.7	(130)	27.7	(67)	7.583	.006**
Should employers be allowed to fire workers who have the AIDS virus?	9.3	(31)	7.4	(18)	0.647	.421
Should the legislature develop legislation to explicitly prohibit discrimination against persons with AIDS in North Carolina?	86.5	(288)	91.3	(221)	3.226	.072
Insurance						
Should health insurance companies be allowed to ask questions to determine if someone belongs to a high-risk group for AIDS before issuing a policy?	47.0	(157)	36.4	(87)	6.407	.011*
Should insurance companies be allowed to exclude coverage specifically for conditions resulting from the AIDS virus?	15.0	(50)	4.2	(10)	17.635	.000***

Note. *df* = degrees of freedom.
* = $P < .05$, ** = $P < .01$, *** = $P < .001$.

At the start of the course, 18% of the students indicated that their blood had been tested for HIV (Table 2); this was less than NHIS cohort findings (26% among 18- to 29-year-olds). Over the period of the course, an additional 5% were tested. Most students reported that HIV testing had occurred through a blood donation, with fewer than a third specifically seeking a blood test voluntarily. When presented with a list of risk factors before the course, 6% of the students identified themselves as members of a high-risk group. This was twice the rate reported for the 18- to 29-year-old cohort in the NHIS. At the end of the course, an additional 3% (amounting to a postcourse total of 9.5%) rated themselves as having behaviors that might place them at risk for HIV infection.

Attitudes toward the rights of health workers to refuse to care for persons infected with HIV changed over the course period (Table 3). There was a significant reduction in the percentage of those who indicated that physicians (–13.7%) and nurses (–13.5%) had such a right. Attitude changes were found, demonstrating by the course end increased understanding and tolerance for persons who are HIV positive.

Anonymous course evaluation forms were distributed after the final class sessions. Students largely rated the course and the lectures as good or excellent. The class most highly ranked was "The Human Side of AIDS," a panel of persons with AIDS. Ninety-five percent of the students would have recommended the course to others.

Discussion

This initiative in university interdisciplinary course development shows that large numbers of university students will

enroll in a credit-bearing course on AIDS. Variable-credit course options will attract undergraduate lecture course students as well as graduate seminar students.

This project indicates that many and significant changes occur in student knowledge and attitudes about AIDS over the period of a university course on AIDS. It suggests that other universities might use this interdisciplinary educational approach and public health initiative to educate students—our future community leaders and professionals—about AIDS. Health sciences professional schools are in a unique position to begin similar campuswide educational programs that demonstrate in education the interdisciplinary nature of the work required in practice. □

Acknowledgments

This research and educational initiative was supported by a grant from the Pew–Rockefeller Health of the Public Program, the UNC Center for Health Promotion and Disease Prevention. This paper was presented in part at the Sixth International Conference on AIDS on June 21, 1990, in San Francisco, CA. It was also partially included in a presentation at the annual meeting of the Association of American Medical Colleges on October 21, 1990, in San Francisco.

Special appreciation goes to Dr. H. Garland Hershey (vice chancellor of Health Affairs and vice provost, University of North Carolina at Chapel Hill) and the Lupton Opportunity Fund for their support of this project.

References

1. Johnson JA, Campbell AE, Toewe CH II, Bell BJ. Knowledge and attitudes about AIDS among first- and second-year medical students. *AIDS Educ Prev.* 1990;2:48–57.
2. Goldman JD. An elective seminar to teach first-year students the social and medical aspects of AIDS. *J Med Educ.* 1987;62:557–561.
3. Hale PJ. An elective course on AIDS. *J Nurs Educ.* 1989;28:232–234.
4. Dommeyer CJ, Marquard JL, Gibson JE, Taylor RL. The effectiveness of an AIDS education campaign on a college campus. *J Am Coll Health.* 1989;38:131–135.
5. Strunin L, Hingson R. Acquired immunodeficiency syndrome and adolescents: knowledge, beliefs, attitudes, and behaviors. *Pediatrics.* 1987;79:825–828.
6. Hingson R, Strunin L, Berlin B. Acquired immunodeficiency syndrome transmission: changes in knowledge and behaviors among teenagers, Massachusetts statewide surveys, 1986 to 1988. *Pediatrics.* 1990;85:24–29.
7. Koopman C, Rotheram-Borus MJ, Henderson R, Bradley JS, Hunter J. Assessment of knowledge of AIDS and beliefs about AIDS prevention among adolescents. *AIDS Educ Prev.* 1990;2:58–70.
8. Royse D, Dhooper SS, Hatch LR. Undergraduate and graduate students' attitudes towards AIDS. *Psychol Rep.* 1987;60:1185–1186.
9. Royse D, Birge B. Homophobia and attitudes toward AIDS patients among medical, nursing, and paramedical students. *Psychol Rep.* 1987;61:867–870.
10. Katzman EM, Mulholland M, Sutherland EM. College students and AIDS: a preliminary survey of knowledge, attitudes, and behavior. *J Am Coll Health.* 1988;37:127–130.
11. Lester LB, Beard BJ. Nursing students' attitudes toward AIDS. *J Nurs Educ.* 1988;27:399–404.
12. Strader MK, Beaman ML. College students' knowledge about AIDS and attitudes toward condom use. *Public Health Nurs.* 1989;6:62–66.
13. Thomas SB, Gilliam AG, Iwrey CG. Knowledge about AIDS and reported risk behaviors among Black college students. *J Am Coll Health.* 1989;38:61–66.
14. Kinnick BC, Smart DW, Bell DA, Blank WR, Gray TR, Schober JL. An assessment of AIDS-related knowledge, attitudes, and behaviors among selected college and university students. *AIDS Public Policy J.* 1989;4:112–119.
15. Gaines J, Iglar AF, Michal ML, Patton RD. Attitudes toward AIDS. *Health Values.* 1988;12:53–60.
16. Crawford J. Attitudes of undergraduate students towards AIDS. *Psychol Rep.* 1990;66:11–16.
17. Hernandez JT, Smith FJ. Inconsistencies and misperceptions putting college students at risk of HIV infection. *J Adolesc Health Care.* 1990;11:295–297.
18. Lesnick H, Pace B. Knowledge of AIDS risk factors in South Bronx minority college students. *J AIDS.* 1990;3:173–176.
19. MacDonald NE, Wells GA, Fisher WA, et al. High-risk STD/HIV behavior among college students. *JAMA.* 1990;263:3155–3159.
20. Thurman QC, Franklin KM. AIDS and college health: knowledge, threat, and prevention at a northeastern university. *J Am Coll Health.* 1990;38:179–184.
21. Carroll L. Concern with AIDS and the sexual behavior of college students. *J Marriage Fam.* 1988;50:405–411.
22. Sy FS, Freeze-McElwee Y, Garrison CZ, Jackson KL. Knowledge, perceived risk, and beliefs about AIDS among high school and college students in South Carolina. *JSC Med Assoc.* 1989;85:481–487.
23. Kelly JA, St Lawrence JS, Smith S Jr, Hood HV, Cook DJ. Medical students' attitudes toward AIDS and homosexual patients. *J Med Educ.* 1987;62:549–556.
24. Manning D, Balson PM, Barenberg N, Moore TM. Susceptibility to AIDS: what college students do and don't believe. *J Am Coll Health.* 1989;38:67–73.
25. Kegeles SM, Adler NE, Irwin CE. Sexually active adolescents and condoms: changes over one year in knowledge, attitudes and use. *Am J Public Health.* 1988;78:460–461.
26. Keeling RP. *AIDS on the College Campus.* Rockville, Md: American College Health Association; 1986. Special Report.
27. Health Resources and Services Administration. Proceedings of the Multidisciplinary Curriculum Development Conference on HIV Infection. Public Health Service; Department of Health and Human Services publication HRP0907199. Washington, DC. 1987.
28. Wertz DC, Sorenson JR, Liebling L, Kessler L, Heeren TC. Knowledge and attitudes of AIDS health care providers before and after education programs. *Public Health Rep.* 1987;102:248–254.
29. Miller TE, Booraem C, Flowers JV, Iversen AE. Changes in knowledge, attitudes, and behavior as a result of a community-based AIDS prevention program. *AIDS Educ Prev.* 1990;2:12–23.
30. Lawrence SA, Lawrence RM. Knowledge and attitudes about acquired immunodeficiency syndrome in nursing and nonnursing groups. *J Prof Nurs.* 1989;5:92–101.
31. Bell RA, Feldman TB, Grissom S, et al. Evaluating the outcomes of AIDS education. *AIDS Educ Prev.* 1990;2:70–83.
32. Dawson DA. AIDS knowledge and attitudes for January–March 1989, provisional data from the National Health Interview Survey (advance data). *Vital and Health Statistics.* National Center for Health Statistics; 1989;176:1–12.

Ineffectiveness of AIDS Education and HIV Antibody Testing in Reducing High-Risk Behaviors among Injection Drug Users

Donald A. Calsyn, PhD, Andrew J. Saxon, MD, George Freeman, Jr, MA, and Stephen Whittaker, ARNP

ABSTRACT

The effectiveness of education in reducing high-risk human immunodeficiency virus (HIV) transmission behaviors was examined in 313 injection drug users. Involvement in high-risk behaviors was assessed via structured interview at study entry and 4 months following the intervention. Subjects were randomly assigned to (1) AIDS education, (2) AIDS education with optional HIV antibody testing, or (3) a wait list. The sample as a whole decreased its involvement in high-risk behaviors, but there were no significant differences as a function of experimental group assignment. (*Am J Public Health.* 1992;82:573–575)

Introduction

Several reports indicate that injection drug users (IDUs) are decreasing their sharing of injection equipment and increasing their use of bleach to clean shared equipment.[1–3] Smaller decreases in high-risk sexual behaviors by IDUs have been found by some but not all investigators.[4] The primary acquired immunodeficiency syndrome (AIDS) prevention tools used to date with IDUs have included mass media education, street outreach education (often involving the provision of bleach and free condoms), human immunodeficiency virus (HIV) antibody testing and counseling, and educational programs for targeted subgroups. Although reductions in risk behaviors are often attributed to these educational and exhortative efforts,[3] no research using randomly assigned control groups has demonstrated the efficacy of specific educational prevention interventions.

The research reported here tested the hypothesis that a 90-minute AIDS education package and knowledge of one's HIV status will reduce high-risk behaviors among IDUs in treatment. A 90-minute educational format was chosen because it seems typical of interventions provided by drug abuse treatment programs with limited resources.

Methods

Three hundred thirteen IDUs receiving or seeking treatment in Seattle, Wash, served as subjects. The subjects were recruited from Evergreen Treatment Services, a nonprofit methadone maintenance program, and the Addictions Treatment Center of the Veterans Affairs Medical Center. The subjects volunteered to participate in an AIDS prevention project and gave informed written consent. They received $30 to defray expenses for each interview. The subjects were representative of clients in treatment in these settings with respect to demographics and involvement in high-risk behaviors.[5]

At enrollment into the project, all subjects were assessed via a structured interview regarding high-risk behaviors.[5] Immediately following this initial interview, all subjects were randomly assigned to one of three AIDS prevention intervention conditions: (1) a 90-minute group AIDS education session, (2) the AIDS education session plus optional HIV antibody testing, or (3) a 4-month wait list.

The AIDS education sessions were conducted by two trained facilitators who lectured from a scripted format while also responding to questions and allowing discussion. The format covered considerable information about the transmission of HIV, the medical aspects of AIDS, HIV antibody testing, and a three-tiered risk reduction strategy: (1) abstinence from illicit drug use, especially injectables, and celibacy; (2) no sharing of injection equipment and long-term monogamy with a known low-risk partner; (3) cleaning of injection equipment with bleach and use of condoms for all sexual activity. A highly emotional videotape, *Drugs and AIDS: An Appeal to Users*,[6] featuring a recovering addict reinforcing the risk reduction strategy, was shown. Cleaning of a syringe with bleach and application of a condom to a dildo were demonstrated. Small vials containing bleach and a starter pack of condoms were given out at the completion of the session.

Four months after the intervention, all subjects who could be located (n = 218) were assessed for behavioral changes in a

Donald A. Calsyn and Andrew J. Saxon are with the Veterans Affairs Medical Center, Seattle, Wash, and the Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle. George Freeman, Jr, is with the People of Color Against AIDS Network, Seattle, Wash, and Southern Illinois University, Carbondale.

Requests for reprints should be sent to Donald A. Calsyn, PhD, Veterans Affairs Medical Center (116DDTP), 1660 South Columbian Way, Seattle, WA 98108.

This paper was submitted to the Journal March 4, 1991, and accepted with revisions August 22, 1991.

TABLE 1—Demographics and High-Risk Behaviors of the Sample

	Initial Assessment (n = 313)		Follow-up Assessment: Completers (n = 218)		Follow-up Assessment: Noncompleters (n = 95)		Analysis[a]
	Mean (SD)	Median	Mean (SD)	Median	Mean (SD)	Median	
Age	38.5 (7.3)	38.0	39.1 (7.8)	37.0	37.1 (5.8)	37.0	$t = 2.5$*
Education	12.8 (2.2)	13.0	12.7 (2.2)	13.0	13.2 (2.0)	13.0	$t = 1.8$
Months in current treatment	20.3 (31.7)	6.0	23.8 (33.8)	8.0	12.4 (24.4)	2.0	$z = 5.3$***[b]
No needle-sharing partners							
Past year	4.1 (11.3)	1.0	3.1 (6.6)	1.0	6.5 (17.6)	1.0	$z = 1.4$
Past 30 days	0.4 (1.4)	0.0	0.3 (0.7)	0.0	0.6 (2.2)	0.0	$z = 0.6$
Males only							
No. female partners	4.0 (15.0)	1.0	2.5 (3.0)	1.0	7.0 (25.1)	2.0	$z = 2.2$*
Condom use, %	9.7 (22.6)	0.0	8.9 (22.9)	0.0	11.0 (22.0)	0.0	$z = 1.4$
Females only							
No. male partners	3.8 (20.5)	1.0	4.1 (22.8)	1.0	2.8 (4.9)	1.0	$z = 0.7$
Condom use, %	12.8 (26.1)	0.0	11.5 (26.3)	0.0	18.9 (25.4)	2.0	$z = 2.2$*
	n (%)		n (%)		n (%)		
Sex							
Male	225 (71.9)		148 (67.9)		77 (81.1)		
Female	88 (28.1)		70 (32.1)		18 (18.9)		$\chi^2 = 5.7$**
Race							
White	215 (68.7)		155 (71.1)		60 (63.1)		
Black	76 (24.3)		50 (22.9)		26 (27.4)		$\chi^2 = 0.3$
Hispanic	10 (3.2)		6 (2.8)		4 (4.2)		
Other	12 (3.8)		7 (3.2)		5 (5.3)		
Marital status							
Married	65 (28.4)		49 (22.5)		16 (16.8)		
Not married	248 (71.6)		169 (77.5)		79 (83.2)		$\chi^2 = 0.3$
Employment status							
Employed	89 (28.4)		60 (27.5)		29 (30.5)		
Unemployed/disabled	224 (71.6)		158 (72.5)		66 (69.5)		$\chi^2 = 0.6$
Primary drug abused							
Opiates	231 (74.5)		177 (81.2)		54 (56.8)		
Cocaine	36 (11.6)		16 (7.3)		20 (21.1)		$\chi^2 = 16.1$***
Other	43 (13.9)		25 (11.5)		21 (22.1)		
Recruitment source							
Methadone program	229 (75.1)		176 (80.8)		53 (58.2)		
Intensive outpatient[c]	27 (8.9)		20 (9.2)		7 (7.7)		
Inpatient[c]	49 (16.1)		18 (8.3)		31 (34.1)		$\chi^2 = 19.7$***
Disinfection of needles with bleach							
Always/usually	35 (17.0)		26 (18.6)		9 (13.6)		
Sometimes/rarely/ never	171 (83.0)		114 (81.4)		57 (86.4)		$\chi^2 = 0.4$

[a]Analyses are based on a comparisons between follow-up completers and noncompleters.
[b]The z estimate is based on the Mann-Whitney U statistic.
[c]Groups were collapsed for statistical analysis.
*$P < .05$
**$P < .01$
***$P < .001$

structured interview similar to the initial interview. The interviewers were blinded to the subjects' educational group assignments. All outcome measures of interest were dichotomized. Changes from initial to follow-up assessment were analyzed by means of the McNemar statistic.

Results

Table 1 contains demographic data for the entire sample and comparisons of follow-up completers and noncompleters. As one might anticipate, completers were more stable than noncompleters. The sample was knowledgeable about AIDS transmission at entry. In response to the open-ended question "What do you know about how people get AIDS?" 90.1% spontaneously mentioned needle sharing and 69.6% mentioned heterosexual activity. To true-false questions concerning means of contracting AIDS, 99% correctly identified needle sharing, 98.7% indicated sexual activity, and 96.8% knew that condoms could reduce risk.

Only 3 subjects (1.4%) of the 213 tested were HIV seropositive.

Table 2 presents the needle-use and sexual risk variables for the sample, initially and at follow-up. On all needle-use variables, significant improvement occurred. Both men and women reduced their numbers of sexual partners; an increase in condom use occurred only for males with multiple partners. No significant differences existed among the three experimental groups on any demographic or high-risk behavior variables at initial assessment. Nor were differences found among experimental groups on any outcome measures.

Discussion

Like other investigators, we observed that IDUs, as a group, were reducing their involvement in high-risk behav-

TABLE 2—Changes in High-Risk Behaviors between Initial and Follow-up Assessments

	Initial Assessment		Follow-up Assessment		
	n	(%)	n	(%)	McNemar
Needle use risk					
Using IV drugs	178	(81.6)	114	(52.3)	$\chi^2 = 47.3$***
Sharing needles[a]	140	(78.7)	54	(46.6)	$\chi^2 = 24.8$***
Cleaning with bleach[b]	7	(13.5)	24	(46.2)	Bimodal $P = .0015$
Sexual behaviors—males					
Multiple partners	65	(43.9)	43	(29.1)	$\chi^2 = 30.6$***
Condom use >50%[c]	0	(0.0)	10	(20.0)	Bimodal $P = .002$
Sexual behaviors—females					
Multiple partners	22	(31.8)	6	(8.7)	Bimodal $P < .001$
Condum use >50%[c]	4	(22.0)	3	(16.6)	Bimodal $P = 1.0$

[a]Only subjects using IV drugs during either period were utilized.
[b]Only subjects sharing injection equipment during either period were utilized.
[c]Only subjects with multiple partners during either period were utilized.
***$P < .001$

iors. Nonetheless, in contradiction to our original hypothesis, we found no differential benefit from a 90-minute AIDS education program with a random-assignment control-group design. Our results accord with those of Gibson et al.[7] and McCusker et al.,[8] who found no differences between experimental educational and minimal control-group interventions when random assignment designs were used.

We feel the failure of the interventions probably reflects their essential misdirection. They were designed primarily to provide information, under the assumption that the subjects had little knowledge about AIDS. In fact, the subjects demonstrated remarkably good knowledge prior to the interventions.

Other factors that may have contributed to the absence of educational effects include the following: (1) The generally low rate of needle sharing in this sample[5,9,10] may have provided a basement effect for the number of needle-sharing partners variable. (2) Informal communication about the education between subjects in the education and wait list conditions could have confounded the intervention. Magura et al.[11] demonstrated that peer influences were the most salient predictors of needle sharing. (3) Public service education and outreach campaigns in progress at the time of the study could have imparted to wait-list subjects information similar to that provided in the intervention. (4) The initial assessment interview may have served as a positive intervention by heightening subject awareness in an already knowledgeable group.

Despite these potential confounders, our results call into question the extensive expenditure of resources on AIDS education that is primarily information-giving in nature with already knowledgeable populations. Future investigations should explore alternative interventions that might be more effective.

The lack of differential behavior change in the subjects receiving HIV testing compared to those not tested may be related to the seronegative status of the vast majority tested. Previous research suggests that persons who are seropositive, but not those who are seronegative, change their behavior after learning their status.[12,13] Our findings suggest that HIV testing in low endemic areas is unlikely to result in reductions in high-risk behaviors among IDUs. This implication does not diminish the importance of HIV testing programs in these areas for seroprevalence monitoring and early medical intervention with persons found to be seropositive. □

Acknowledgments

This research was supported by a grant (RO1 DA5281) from the National Institute of Drug Abuse and by the Medical Research Service of the Department of Veterans Affairs.

At the time of the study, Stephen Whittaker was with the Veterans Affairs Medical Center, Seattle, Wash. He has since died.

This paper was presented in part at the Sixth International Conference on AIDS, San Francisco, Calif, June 1990, and at the Annual Meeting of the American Psychological Association, Boston, Mass, August 1990.

References

1. Guydish JR, Abramowitz A, Woods W, Black DM, Sorensen JL. Changes in needle sharing behavior among intravenous drug users: San Francisco, 1986–1988. *Am J Public Health.* 1990; 80:995–997.
2. van den Hoek JAR, van Haastrecht HJA, Coutinho RA. Risk reduction among intravenous drug users in Amsterdam under the influence of AIDS. *Am J Public Health.* 1989;79:1355–1357.
3. Stephens RC, Feucht TE, Roman SW. Effects of an intervention program on AIDS-related drug and needle behavior among intravenous drug users. *Am J Public Health.* 1991;81:568–571.
4. Higgins DL, Galavotti C, Johnson R, O'Reilly KR, Rugg DL. The effect of HIV antibody counseling and testing on risk behaviors: Are the studies consistent? *Sixth International Conference on AIDS;* 1990;1:335. Abstract.
5. Calsyn DA, Saxon AJ, Freeman Jr. G, Whittaker S: Needle-use practices among intravenous drug users in an area where needle purchase is legal. *AIDS.* 1991;5:187–193.
6. 3 Head Films, producers. *Drugs and AIDS: An Appeal to Users* [videotape]. Seattle, Wash: Video Communication Systems Inc; 1987.
7. Gibson DR, Lovelle-Drache J, Young M, Sorensen JL. Does brief counseling reduce risk in IV drug users? Final results from a randomized clinical trial. *Seventh International Conference on AIDS;* 1991;2:70.
8. McCusker J, Stoddard A, Zapka J, et al. Behavioral outcomes of AIDS educational interventions for drug users in treatment. *Seventh International Conference on AIDS;* 1991;2:70.
9. Schoenbaum EE, Hartel D, Selwyn PA, et al. Risk factors for human immunodeficiency virus infection in intravenous drug users. *N Engl J Med.* 1989;321:874–879.
10. Chaisson RE, Bacchetti P, Osmond D, Brodie B, Sande MA, Moss AR. Cocaine use and HIV infection in intravenous drug users in San Francisco. *JAMA.* 1989;261:561–565.
11. Magura S, Grossman JL, Lipton DS, et al. Determinants of needle sharing among intravenous drug users. *Am J Public Health.* 1989;79:459–462.
12. van den Hoek A, van Haastrecht HJA, Coutinho RA. Heterosexual behavior of intravenous drug users in Amsterdam: implications for the AIDS epidemic. *AIDS.* 1990;4:449–553.
13. Cates W, Handsfield H. HIV counseling and testing: does it work? *Am J Public Health.* 1988;78:1533–1534.

Human Immunodeficiency Virus Epidemic. Washington, DC: US Government Printing Office; 1988. Publication 0-214-701:QL3.
3. McCormick K. AIDS: moving toward chronic care. *Med Health Perspect.* 1989;43(suppl 11):1–4.
4. Weinberg DS, Murray HW. Coping with AIDS: the special problems of New York City. *N Engl J Med.* 1987;317:1469–1472.
5. Taravella S. Reserving a place to treat AIDS patients in the hospital. *Mod Healthcare.* 1989;33–37.
6. Strain JJ, Fahs M, Fulop G, Sacks H. AIDS: epidemiology and treatment issues. *Mt Sinai J Med.* 1989;56:233–237.
7. Oberlink M. The AIDS cluster: New York's Mount Sinai establishes inpatient unit. *AIDS Patient Care.* 1988;2:26–29.
8. Centers for Disease Control. Human immunodeficiency virus (HIV) infection classification. *MMWR.* 1987;36(S-7):1–20.
9. Turner BJ, Kelly JV, Ball JK. A severity classification system for AIDS hospitalizations. *Med Care.* 1989;27:423–437.
10. Kelly JV, Ball JK, Turner BJ. Duration and costs of AIDS hospitalizations in New York: variations by severity of illness. *Med Care.* 1989;27:1085–1098.
11. Ball JK, Turner BJ. Distribution and characteristics of AIDS hospitalizations in the US, 1986–1987. Proceedings of the Fifth International Conference on AIDS; June 4–9, 1989; Montreal, Canada.
12. National Center for Health Services Research and Health Care Technology Assessment/Hospital Studies Program. *A Severity Classification System for AIDS Hospitalization.* Springfield, Va: National Technical Information Service; 1988: Publication PB88-249412:D1-1-D1010.

Duration of Medicaid AIDS Hospitalizations—Variation by Season, Stage, and Year

Leona Enama Markson, ScD, Barbara J. Turner, MD, MSEd, and Thomas R. Fanning, PhD

ABSTRACT

Using New York State Medicaid data from 1984 to 1987, we analyzed hospital length-of-stay patterns for acquired immunodeficiency syndrome (AIDS) patients. We found an overall decline in monthly average length of stay, with seasonal fluctuations producing longer stays in the fall and winter months. These findings suggest that hospital utilization for AIDS patients is changing over time and may vary by season. (*Am J Public Health.* 1992;82:578–580)

Introduction

Over the past decade, hospitals in several regions of the country have been heavily burdened by the dramatic rise in the number of acquired immunodeficiency syndrome (AIDS) patients.[1,2] Inpatient care is one of the most expensive components of the cost of AIDS,[3–5] and hospitals in New York State have one of the longest average lengths of stay for AIDS care in the country.[2,6,7] Information on patterns of AIDS hospitalizations is needed to assist hospital planners in meeting the needs of this special population. In this study, we analyzed length-of-stay patterns for New York State Medicaid AIDS hospitalizations to examine temporal patterns and changes in hospital length of stay.

Methods

We examined hospital claims for New York State Medicaid AIDS cases over 35 months, from October 1984 through August 1987. Using a methodology tested elsewhere,[8] we identified a hospital stay as an AIDS admission if one of the following ICD-9-CM codes was recorded: *Pneumocystis carinii* pneumonia (136.3), selected immunodeficiency codes (279.10 and 279.19), or the AIDS codes (042.xx or 043.xx).

The study population comprised AIDS patients aged 15 to 65 years who were hospitalized for nonobstetric reasons. Two classes of outliers were deleted from the sample: patient admissions with stays of more than 66 days (less than 5% of the patients) and those with more than eight inpatient stays over the 35-month period (1.5% of the patients). The study sample consisted of 11 482 patient admissions; males constituted approximately three fourths of the sample. Eighty-two percent of the sample was 26 to 45 years of age at the time of hospitalization.

To adjust for differences in case mix, we determined the severity of each case, using the Staging Classification for AIDS Hospitalizations.[9,10] This AIDS staging system can be summarized into three integer stages, with stage 3 indicating the most severe cases. Stage 4 is used in some analyses to distinguish patients who died in the hospital.

Length of stay was analyzed as an average per month, calculated as the av-

Leona Enama Markson and Barbara J. Turner are with the Center for Research in Medical Education and Health Care and the Department of Medicine, Jefferson Medical College, Philadelphia, Pa. Thomas R. Fanning is with the New York State Department of Social Services, Albany.

Requests for reprints should be sent to Leona Enama Markson, ScD, Center for Research in Medical Education and Health Care, Jefferson Medical College, 1025 Walnut St, Suite 119, Philadelphia, PA 19107-5083.

This paper was submitted to the Journal November 15, 1990, and accepted with revisions November 12, 1991.

erage length of a hospital stay for all patients admitted in the same calendar month. We observed patterns in length of stay by plotting 3-month moving averages, a useful smoothing technique to reduce the short-term fluctuations in time-series data.[11] To compute 3-month moving averages, we determined the average length of stay for all patients admitted from October 1984 through December 1984, then determined the 3-month average for November 1984 through January 1985, and so on. Moving averages were used only in plotting to assist in visual analysis of trends.

To assess the influence of stage, year, and season, we used multiple linear regression to model monthly variation in hospital stays across the total population. The sample size for this analysis equaled 140 "stage-months" of observation (35 months of utilization and discharge data, classified into four stages). For example, the average length of stay for all Stage 1 patients admitted in the first month of the study is one of the 140 observations. Natural logarithms were used to transform average length of stay per month to produce a more normal distribution of the dependent variable.

Results

Figure 1 displays two trends in monthly average length of stay for New York Medicaid AIDS cases: (1) an overall downward trend ($r = -.62$, $P < .001$) and (2) a periodic trend, with longer stays in the fall and winter months. The longest average length of stay was 21.2 days, in October 1984 (month 1 of the study); the shortest was 13.8 days, in August 1987 (month 35). For each of the federal fiscal years (FFYs) of the study, the average length of stay for the month of January was compared with the average length of stay for the month of June; length of stay was longer in January by 2.8 days in FFY 1985, 1.1 days in FFY 1986, and 2.4 days in FFY 1987. The periodic trend in monthly average length of stay remained after patients who died in the hospital were removed from the analysis.

Overall, the average length of stay was 13.0, 19.2, and 20.3 days for patients in stages 1, 2, and 3, respectively. Average length of stay by fiscal year was 17.5 days in FFY 1985, 16.4 days in FFY 1986, and 15.9 days in FFY 1987. The regression analyses shown in Table 1 indicate that stage, year, and season explained 76% of the variance in monthly length-of-stay patterns ($F = 50.5$, $P < .001$). Cases classified in stages 2 and 3 had stays that were 58% longer than the stays of cases in stage 1 ($P < .001$), and hospital stays were on average 7% longer ($P < .05$) in fall and winter months than in summer months. After adjustments were made for stage and fiscal year, hospital stays declined by approximately 6% in FFY 1986 and 11% in FFY 1987 as compared with FFY 1985 ($P < .001$). These numbers represent approximately a 1-day decline in hospital length of stay per year.

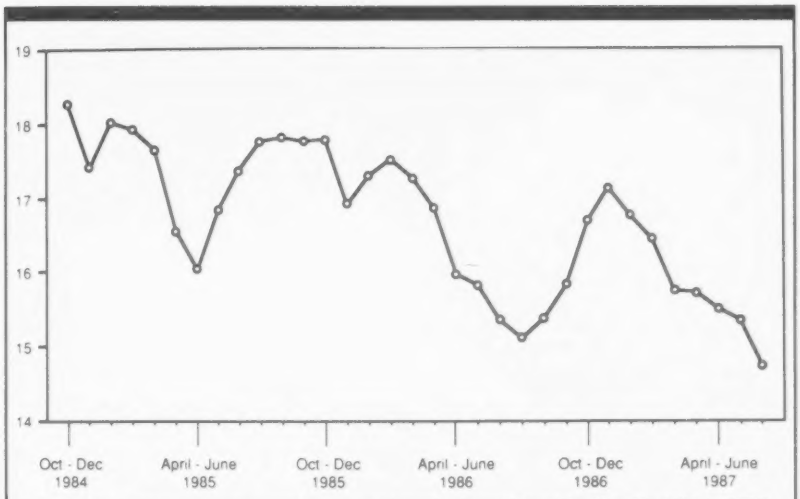


FIGURE 1—Three-month moving average length of stay (days) for New York State AIDS patient admissions from October 1984 through August 1987.

Discussion

Our analyses indicate that length of hospital stay for Medicaid AIDS patients in New York State has declined in recent years, but that average length of stay per month varies depending upon the season of the year and the case mix of the population. Although the effect of season was not large after adjustments were made for severity of illness, seasonal variation in hospital utilization for AIDS patients needs to be explored further. AIDS patients are susceptible to conditions such as respiratory infections, for which prevalence rates and case fatality rates are known to peak in the winter months.[12–14] Other explanations, independent of clinical issues such as availability of nursing home beds, may contribute to seasonal fluctuations in the demand for hospital care. There have been reports on seasonal variations in hospital crowding in New York City in recent times,[15] and the influence of AIDS care in these seasonal hospital occupancy problems should be explored.

This study does not address reasons for the declining length of stay; however, the trend probably reflects improvements in clinical diagnosis and treatments, as well as improved outpatient care. A similar decline in length of hospital stay for a sample of AIDS patients in Massachusetts was recently reported.[16] In addition, data from the National Hospital Discharge Survey for all types of hospitalized patients discharged from short-stay hospitals in the United States indicate an overall decline in recent years in the average length of stay.[17] This national downward trend in length of stay across all diagnosis groups, which is often attributed to prospective payment systems, may also influence the hospital stay patterns for AIDS cases. However, per-case payment for New York Medicaid hospitalizations was not instituted until 1988, after the time of the study.

It should be noted that the observed decline in length-of-stay patterns may not be uniform across all hospitals in New York State. Our analysis focused on monthly variation in length-of-stay patterns and not on length-of-stay patterns within individual hospitals. In addition, our sample was restricted to Medicaid AIDS patients, and it is not clear whether these trends hold for patients not covered by Medicaid. In recent studies of hospital services for AIDS patients, Medicaid was found to be the most frequent payer of AIDS inpatient care.[16,18] Moreover, the number of AIDS patients enrolled in Medicaid is on the rise.[19–21]. As a result, patterns of inpatient care for Medicaid AIDS

TABLE 1—Regression Model of Monthly Log Average Length of Stay (October 1984 through August 1987)

Variable	Coefficient	Standard Error	*P* Value
Stage 2	0.46	.028	<.001
Stage 3	0.46	.028	<.001
Stage 4	0.40	.028	<.001
FFY 1986	−0.07	.024	<.001
FFY 1987	−0.12	.025	<.001
Fall	0.07	.029	.01
Winter	0.07	.028	.02
Spring	0.02	.029	.43
Intercept	2.55	.030	<.001

Note. The sample size for this analysis equaled 140 "stage-months" of observation (35 months of utilization and discharge data, classified into four stages). The reference groups were Stage 1 (the least severe cases), the summer season (June, July, and August), and federal fiscal year (FFY) 1985. $R^2 = .76$; Adj. $R^2 = .74$; $F = 50.5$, $P < .001$.

cases can influence the availability of hospital care for all types of patients. □

Acknowledgments

This research was supported in part by the Health Care Financing Administration Cooperative Agreement 18C-9924219-01.

An earlier version of this paper was presented by Barbara J. Turner at the Fifth International Conference on AIDS, June 5, 1989, in Montreal, Canada.

The authors wish to thank Linda McKee for assistance in data processing and Patricia Lovec for assistance in manuscript preparation.

References

1. Andrulis DP, Beers VS, Bentley JD, Gage LS. The provision and financing of medical care for AIDS patients in U.S. public and private teaching hospitals. *JAMA.* 1987;258:1343–1346.
2. Green J, Singer M, Wintfeld N, Schulman K, Passman L. Projecting the impact of AIDS in hospitals. *Health Aff (Millwood).* 1987;6:19–31.
3. Sisk J. The cost of AIDS: a review of estimates. *Health Aff (Millwood).* 1987;6:5–21.
4. Scitovsky AA, Rice DP. Estimates of the direct and indirect costs of acquired immunodeficiency syndrome in the United States. *Public Health Rep.* 1987;102:5–17.
5. Scitovsky AA. The economic impact of AIDS in the United States. *Health Aff (Millwood).* 1988;7:32–45.
6. Ron A, Rogers DE. New York City's health care crisis: AIDS, the poor, and limited resources. *JAMA.* 1988;260:1453.
7. Ball JK, Turner BJ. *AIDS in U.S. Hospitals, 1986–1987: A National Perspective.* Washington, DC: US Dept of Health and Human Services; August 1991. Hospital Studies Program Research Note 15, publication AHCPR 91-0015.
8. Keyes M, Andrews R, Mason ML. A methodology for building an AIDS research file using Medicaid claims and administrative data bases. *J AIDS.* 1991;4:1015–1024.
9. Turner BJ, Kelly JV, Ball JK. A severity classification system for AIDS hospitalizations. *Med Care.* 1989;27:423–437.
10. Kelly JV, Ball JK, Turner BJ. Duration and costs of AIDS hospitalizations in New York: variations by patient severity of illness and hospital type. *Med Care.* 1989;27:1085–1098.
11. Pindyck RS, Rubinfeld DL. *Economic Models and Economic Forecasts.* 2nd ed. New York, NY: McGraw-Hill; 1981.
12. Leads from the MMWR: increase in pneumonia mortality among young adults and the HIV epidemic—New York City, United States. *JAMA.* 1988;260:2181–2185.
13. Increase in pneumonia mortality among young adults and the HIV epidemic—New York City, United States. *MMWR.* 1988;37:593–596.
14. Flournoy DJ, Stalling FH, Catron TL. Seasonal and monthly variation of and other pathogens in bacteremia (1961–1981). *Ecology of Disease.* 1983;2(2):157–160.
15. Barbanel J. After crisis, unused beds increase in hospitals across New York City. *New York Times.* July 7, 1990; A1.
16. Seage GR, Landers S, Lamb GA, Epstein AM. Effect of changing patterns of care and duration of survival on the cost of treating the acquired immunodeficiency syndrome (AIDS). *Am J Public Health.* 1990;80:835–839.
17. Pokras R, Kozak LJ, McCarthy E, Graves EJ. Trends in hospital utilization, 1965–86. *Am J Public Health.* 1990;80:488–490.
18. Andrulis DP, Weslowski VB, Gage LS. The 1987 U.S. hospital AIDS survey. *JAMA.* 1989;262:784–794.
19. Kizer KW, Keith J, Hiehle G, Wolcott C, Maxfield WT. *Expenditures, demographics and mortality for persons with AIDS on MediCal.* Sacramento, Calif: California Department of Health Services; 1988.
20. Roper WL, Winkenwerder W. Making fair decisions about financing care for persons with AIDS. *Public Health Rep.* 1988;103:305–308.
21. Green J, Arno PS. The 'Medicaidization' of AIDS, trends in the financing of HIV-related medical care. *JAMA.* 1990;24:1261–1266.

Decline in Human Immunodeficiency Virus Seropositivity and Seroconversion in US Navy Enlisted Personnel: 1986 to 1989

Frank C. Garland, PhD, Edward D. Gorham, MPH, Steven O. Cunnion, CAPT, MC, USN, Milan R. Miller, Louis L. Balazs, MS, and the Navy HIV Working Group

ABSTRACT

The US Navy administered 1 795 578 enzyme-linked immunosorbent assay (ELISA) tests to 848 632 active-duty Navy enlisted personnel during 1986 to 1989. This study identified 2438 human immunodeficiency virus (HIV)-seropositive active-duty enlisted Navy personnel, including 778 seroconverters. Three types of quarterly rates of HIV seropositivity and seroconversion were determined. All three rates declined. This decline could not be explained by changes in the population tested according to age, race, sex, occupation, or geographic location of home port. (*Am J Public Health.* 1992;82:581–584)

Introduction

The course of the human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome epidemic has been reported in various populations, including homosexual and bisexual men,[1–5] prisoners,[6–8] and military recruits.[9–11] This study reports quarterly time-trends in HIV seropositivity and seroconversion rates among active-duty enlisted US Navy personnel.

HIV seroconversion rates were previously reported for 1986 to 1988 for the Navy (0.70 per 1000 person-years) and Marine Corps (0.27 per 1000 person-years).[12] Army rates (0.77 per 1000 person-years)[13] were similar to those reported for the Navy.

Methods

The Naval Health Research Center in San Diego, Calif, maintains computerized career records for all Navy personnel and HIV testing histories for all those receiving enzyme-linked immunosorbent assay (ELISA) and Western blot tests. Many individuals had several tests during 1986 to 1989 as a result of routine screening and testing of personnel deployable overseas or health care providers.[12] Retesting allowed identification of seroconverters (i.e., persons with an initial negative test followed by a positive test).

All seropositive individuals had two positive paired ELISA tests (from the same blood collection) followed by at least one positive Western blot assay. The criterion for a positive Western blot assay was a specimen exhibiting at least two of three bands at p24, gp41, or gp120/160.[12] Sources of data and testing procedures have been described in detail elsewhere.[12]

For completeness in assessing trends, three rates were calculated: (1) the rate of newly identified HIV-seropositive individuals per 1000 tested, (2) the rate of HIV seroconversion per 1000 individuals tested with a previous negative test, and (3) the rate of HIV seroconversion per 1000 person-years. Person-years were computed as the time between an individual's first negative test and his or her first positive or last negative test. Persons with two negative tests were assumed to be negative through the end of the study period.

Crude rates based on person-years were dependent upon the number of tests administered during an interval, which varied among quarters. Therefore, these rates were adjusted for the number of tests per quarter.[14] Indirect age-adjustment (using age-specific Navy rates for the entire study period as the standard) was performed for the rate of HIV seroconversion per 1000 individuals tested with a previous negative test.[15] To determine if the population tested varied over time by age, race, sex, occupation, or geographic location, the percentage distribution of each of these factors by quarter was determined.

Frank C. Garland, Edward D. Gorham, Milan R. Miller, and Louis L. Balazs are with the Division of Epidemiology, Department of Health Sciences and Epidemiology, Naval Health Research Center, San Diego, Calif. Frank C. Garland is also with the Department of Community and Family Medicine, University of California, San Diego. Steven O. Cunnion is with the United States Uniformed Services University of the Health Sciences, Bethesda, Md.

Requests for reprints should be sent to Frank C. Garland, PhD, Division of Epidemiology, Naval Health Research Center, PO Box 85122, San Diego, CA 92186-5122.

This paper was submitted to the Journal June 10, 1991, and accepted with revisions September 16, 1991.

Note. This report was supported by the Bureau of Medicine and Surgery. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Navy, the Department of Defense, or the US government. No endorsement by the Department of the Navy has been given.

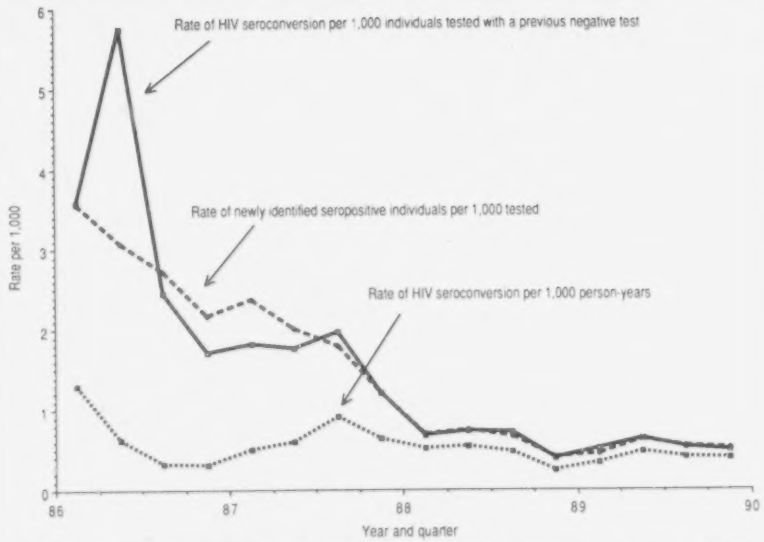


FIGURE 1—HIV seropositive and seroconversion rates, by quarter, active-duty US Navy enlisted personnel, 1986–1989.

Results

A total of 1 795 578 ELISA tests were administered to 848 632 individuals (mean number of tests per quarter, 112 224), and 2438 HIV-seropositive individuals were identified, of which 778 were seroconverters. Ninety-nine percent of the active-duty Navy enlisted population received ELISA tests during the study period.

The rate of newly identified seropositive individuals per 1000 tested declined from 3.6 in the first quarter of 1986 to 0.5 in the last quarter of 1989 ($P < .05$) (Figure 1). The mean annual rate declined from 2.9 in 1986 to 0.5 in 1989 ($P < .05$) (Table 1).

The rate of HIV seroconversion per 1000 individuals tested with a previous negative test peaked at 5.7 during the second quarter of 1986, then declined to 2.5 in the next quarter, a nonsignificant difference (Figure 1). The rates plateaued until the third quarter of 1987. This was followed by a decrease to a relatively stable level throughout 1988 and 1989, reaching 0.5 per 1000 in the fourth quarter of 1989, a lower HIV seroconversion rate ($P < .05$) than all quarterly rates in 1986 and 1987. The mean annual rate of HIV seroconversion per 1000 individuals tested with a previous negative test declined from 3.4 in 1986 to 0.5 in 1989 ($P < .05$) (Table 1).

There was little quarterly variation in the rate of HIV seroconversion per 1000 person-years (Figure 1). The mean annual rate in 1986, 0.9 per 1000 person-years, was more than two times higher ($P < .05$) than the mean annual rate of 0.4 in 1989 (Table 1).

The mean annual rate of HIV seroconversion per 1000 individuals tested with a previous negative test was higher for Blacks than for Whites in all 4 years ($P < .05$) (Figure 2). Seroconversion rates in individuals other than White or Black (i.e., Oriental, Filipino, and American Indian) were similar to those of Blacks in the first quarter of the study, but approximated those of Whites thereafter. A steep decline in rates occurred in persons 30 years of age or older between the first and third quarters of 1986 (Figure 3).

The proportion of the population tested by race, sex, and occupational risk group was similar in each quarter (not shown). There were some minor differences according to age.

Discussion

HIV seropositivity and seroconversion rates declined in the Navy during 1986 to 1989. The rate of newly identified seropositive individuals per 1000 tested and the rate of seroconversion per 1000 individuals tested with a previous negative test showed a more marked decline than did the rate of HIV seroconversion per 1000 person-years. The declining pattern in quarterly rates suggests the tail of an epidemic curve,[16,17] which may have peaked in or prior to 1986. The declining trend in the Navy was similar to that reported for the Army[18] and was consistent with the decline in seroprevalence rates observed in military recruits during 1985 to 1989.[11] Rates were markedly higher in Blacks than in Whites and other races in the early study period, which was also consistent with previous reports.[10,12,13]

The absence of a strong trend in the rate of HIV seroconversion per 1000 person-years may be partly explained by the method of calculation, which tended to reduce differences between quarterly rates. Nearly the entire population con-

TABLE 1—Annual Mean Seropositive and Seroconversion Rates per 1000 Tested: Active-Duty US Navy Enlisted Personnel, 1986–1989

Year	Number of newly-identified seropositives	Rate of newly seropositive individuals per 1000 tested	95% Confidence Limits		Number of Seroconverters	Rate of HIV Seroconversion per 1000 Individuals with a Previous Negative Test (Age-Adjusted)	95% Confidence Limits		Rate of HIV Seroconversion per 1000 Person-Years (Adjusted for Number of Tests)	95% Confidence Limits	
			Lower	Upper			Lower	Upper		Lower	Upper
1986	1197	2.9	2.7	3.1	69	3.4	2.7	4.3	0.9	0.7	1.1
1987	602	1.8	1.7	2.0	193	1.7	1.5	1.7	0.7	0.6	0.8
1988	425	0.6	0.5	0.7	328	0.6	0.5	0.7	0.4	0.4	0.4
1989	214	0.5	0.4	0.6	188	0.5	0.4	0.6	0.4	0.3	0.5
1986–1989	2438	1.5	1.5	1.5	778	1.6	1.5	1.7	0.5	0.5	0.5

tributed to each quarterly interval, whether or not the individuals were tested in a particular quarter.

The decline in rates was unlikely to be due to a change in the population tested over time, inasmuch as the population tested did not vary importantly by race, sex, or occupational risk group in each quarter. New York, San Francisco, and Washington, DC, were high-risk areas for HIV seropositivity in military recruits,[9] but the proportion of the population tested from these areas was not overrepresented during the periods of peaks in HIV rates.

There is currently no detailed information regarding trends in the Navy in behavior such as high-risk sexual activities and intravenous drug use, so the reasons for the decrease in HIV rates in the Navy are speculative. National educational campaigns and Navy-sponsored educational programs[19] may have succeeded in reducing some high-risk behavior. To the extent that HIV transmission occurs among active-duty Navy personnel, policies of exclusion of seropositive applicants and testing of active-duty personnel may have decreased the population of HIV carriers (and perhaps seronegative personnel with high-risk behaviors), leading to a reduction in the likelihood of HIV transmission within the Navy.

Most other studies of trends have been limited to high-risk groups,[1–8,20–22] which are presumably different from the Navy population in prevalence of risk factors. Several cautions must be considered regarding the generalization of the trends found in the Navy to other populations. The Navy is not representative of the greater US population in many respects, including its age and sex structure, and it specifically prohibits high-risk behavior.[23–26] The decline reported here in HIV seropositivity rates is encouraging, but may in part reflect a pattern imposed on the Navy population by testing and exclusion of HIV seropositive applicants. □

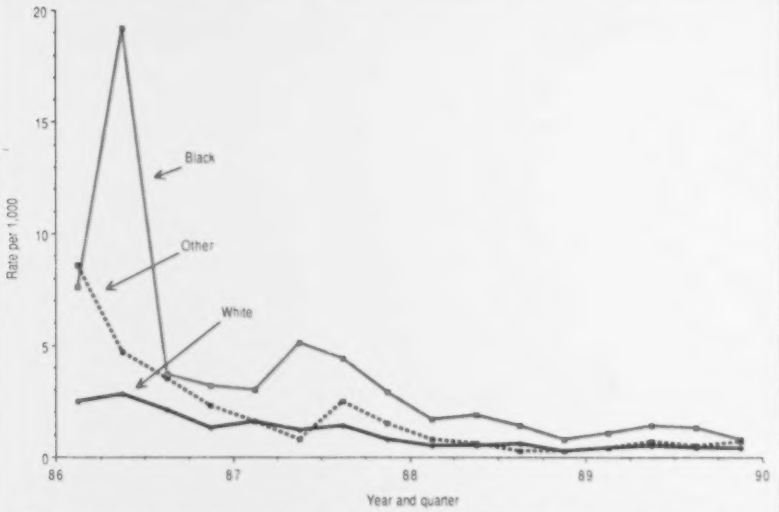


FIGURE 2—Rate of HIV seroconversion per 1000 individuals tested with a previous negative test, adjusted for age, by race, by quarter, active-duty US Navy enlisted personnel, 1986–1989.

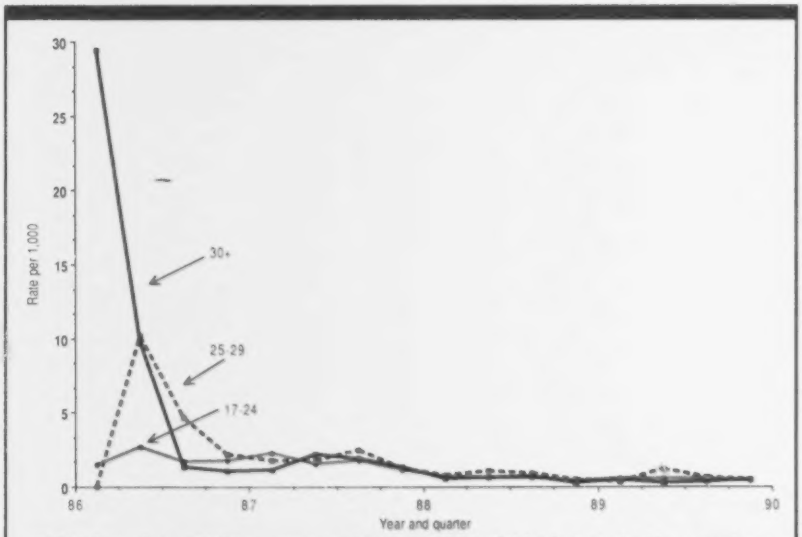


FIGURE 3—Rate of HIV seroconversion per 1000 individuals tested with a previous negative test, by age, by quarter, active-duty US Navy enlisted personnel, 1986–1989.

Acknowledgments

Special thanks are due CDRs T. Hickey and E. Edman; Hospital Corpsman, Chief, Richard Gates; Hospital Corpsman, First Class, Janice Moore; and Hospital Corpsman, Chief, David Jackson and Ms Edith Smith for their diligent work in maintaining records of HIV-tested personnel. CAPT W. Karney provided support for this project. Tammy Clark and Laura Schaffroth assisted in data analysis. Dr E. K. Eric Gunderson brought together many of the data bases that provided the demographic information for this report. Mr Larry R. Bigbee provided assistance with assembly of data bases.

References

1. Winkelstein W Jr, Wiley JA, Padian NS, et al. The San Francisco Men's Health Study: continued decline in HIV seroconversion rates among homosexual/bisexual men. *Am J Public Health.* 1988;78:1472–1474.
2. Hessol NA, Lifson AR, O'Malley PM, et al. Prevalence, incidence, and progression of human immunodeficiency virus infection in homosexual and bisexual men in hepatitis B vaccine trials. *Am J Epidemiol.* 1989;130:1167–1175.
3. Echenberg D, Rutherford G, O'Malley PM, et al. Acquired immunodeficiency syndrome in the San Francisco cohort study, 1978–1985. *MMWR.* 1985;34:1–2.
4. Lemp GF, Payne SF, Rutherford GW, et al. Projections of AIDS morbidity and mortality in San Francisco. *JAMA.* 1990;263:1497–1501.
5. Guggenmoos-Holzmann I, Deininger S, Laukamm-Josten U, et al. Time trends of HIV-I antibody prevalence in homosexual men. *Int J Med Microbiol.* 1990;272:347–354.
6. Vlahov D, Munoz A, Brewer F, et al. Seasonal and annual variation of antibody to HIV-1 among male inmates entering Maryland prisons: update. *AIDS.* 1990;4:345–350.

7. Horsburgh CR, Jarvis JQ, McArthur T, et al. Seroconversion to human immunodeficiency virus in prison inmates. *Am J Public Health.* 1990;80:209–210.
8. Morse DL, Truman BI, Hanrahan JP, et al. AIDS behind bars: epidemiology of New York State prison inmate cases, 1980–1988. *NY State J Med.* 1990;90:133–138.
9. Burke DS, Brundage JF, Herbold JR, et al. Human immunodeficiency virus infections among civilian applicants for United States military service, October 1985 to March 1986: demographic factors associated with seropositivity. *N Engl J Med.* 1987;317: 131–136.
10. Burke DS, Brundage JF, Goldenbaum M, et al. Human immunodeficiency virus infections in teenagers: seroprevalence among applicants for U.S. military service. *JAMA.* 1990;263:2074–2077.
11. Brundage JF, Burke DS, Gardner LI, et al. Tracking the spread of the HIV infection epidemic among young adults in the United States: results of the first four years of screening among civilian applicants for military service. *J AIDS.* 1990;3:1168–1180.
12. Garland FC, Douglas ML, Hickey TM, et al. Incidence of human immunodeficiency virus seroconversion in US Navy and Marine Corps personnel, 1986 through 1988. *JAMA.* 1989;262:3161–3165.
13. McNeil JG, Brundage JF, Wann ZF, et al. Direct measurement of human immunodeficiency virus seroconversion in a serially tested population of young adults in the United States Army, October 1985 to October 1987. *N Engl J Med.* 1989;320:1581–1585.
14. Kramer MS. *Clinical Epidemiology and Biostatistics.* New York, NY: Springer-Verlag; 1988:32–35.
15. Lilienfeld AM, Lilienfeld DE. Foundations of epidemiology. 2nd ed. New York, NY: Oxford University Press; 1980:335–337.
16. Langmuir A. AIDS projections are too high. *Pan Am Health Org Bull.* 1989;23: 121–126.
17. Bregman DJ, Langmuir AD. Farr's law applied to AIDS projections. *JAMA.* 1990;263:1522–1525.
18. McNeil JG, Brundage JF, Gardner LI, et al. Trends of HIV seroconversion among young adults in the US Army, 1985 to 1989. *JAMA.* 1991;265:1709–1714.
19. McGrane WL, Toth FJ. The use of interactive media for HIV/AIDS prevention in the military community. *Milit Med.* 1990;6:235–240.
20. Hahn RA, Onorato IM, Jones TS, et al. Prevalence of HIV infection among intravenous drug users in the United States. *JAMA.* 1989;261:2677–2684.
21. Schoenbaum EE, Hartel D, Selwyn PA, et al. Risk factors for human immunodeficiency virus infection in intravenous drug users. *N Engl J Med.* 1989;321:874–879.
22. Chaisson RE, Moss AR, Onishi R, et al. Human immunodeficiency virus infection in heterosexual drug users in San Francisco. *Am J Public Health.* 1987;77:169–172.
23. Herbold JR. AIDS policy development within the Department of Defense. *Milit Med.* 1986;151:623–627.
24. Hoffman KJ. 1989 Sir Henry Wellcome medal and prize winner: mandatory HIV testing and the duty to screen. *Milit Med.* 1990;155:195–201.
25. Tramont EC, Redfield R, Burke D, et al. HTLV-III/LAV infections in the military. *Milit Med.* 1987;152:105–106.
26. Mayer W. What we are doing about AIDS. *Defense.* 1986;1:30–32.

Fear of Dying and HIV Infection vs Hepatitis B Infection

Lawrence J. Schneiderman, MD, and Robert M. Kaplan, PhD

ABSTRACT

Accidental exposure to the blood of hepatitis B patients produced less fear than does accidental exposure today to the human immunodeficiency virus (HIV), even though both have an approximately equal overall risk of death (~1%).

Subjects responding to hypothetical insect-exposure and disease-exposure scenarios chose to avoid the HIV-type risk of 1% chance of exposure/100% chance of death.

Fear of certain death seems to account for the greater concern about exposure to HIV than to Hepatitis B. (*Am J Public Health.* 1992;82:584–586)

Introduction

Health care providers are at risk of infection resulting from accidental exposure to blood, such as needlestick injury. For many years before a vaccine became available, providers risked infection and death from hepatitis B. These risks were faced without the outcry and debate and the disruption of services that now attend the acquired immunodeficiency syndrome (AIDS) crisis. One of us, in a lecture to health care providers, pointed out this apparent irrationality of health providers' behavior in response to human immunodeficiency virus (HIV). As can be demonstrated by simple calculations, the probability of death from accidental exposure to HIV is no worse than was the probability of death from accidental exposure to hepatitis B in the recent past. The chance of hepatitis B infection following accidental exposure is in the range of 25%; we conservatively assume that about 5% of those infected die either from fulminant or chronic progressive disease.[1,2] The chance of HIV infection following accidental exposure is much lower, apparently no greater than 1%, and it is presently assumed that 100% of patients who are HIV infected will develop AIDS and its complications and die.[3]

Multiplying the probabilities of risk of infection following exposure by the risk of death following infection yields approximately equal outcomes. That is, in both circumstances there is an approximately 1% chance of dying following accidental exposure. Why then are health providers, who took relatively few precautions in the hepatitis B era, now taking exceptional precautions, some of them even refusing

The authors are with the Division of Health Care Sciences, Department of Community & Family Medicine, University of California, San Diego.

Requests for reprints should be sent to Lawrence J. Schneiderman, MD, Division of Health Care Sciences, 0622, Department of Community & Family Medicine, University of California, San Diego, La Jolla, CA 92093-0622.

This paper was submitted to the Journal March 25, 1991, and accepted with revisions July 9, 1991.

to care for patients who are HIV infected? Are they sharing in society's "AIDS hysteria," much of which is known to be fanned by prejudice and ignorance? A member of the audience provided a thought-provoking answer: "The problem is that once you have HIV infection you have no chance. You know you're going to die." The answer suggested that even though the probability of infection is very low, the certainty of death might be so undesirable that it might account for the greater concern about HIV than about hepatitis B—a hypothesis we chose to test in an empirical study.

Expected utility theory[4,5] has traditionally served to explain decision-making processes that involve risk. According to this theory, individuals should favor the alternative that maximizes gains and minimizes losses. In the case of HIV and hepatitis B exposure, the expected utility of either alternative is nearly equal. Thus, it would be expected that the motivation to avoid each condition is approximately equal.

An alternative to expected utility theory was introduced more recently. This theory, known as prospect theory,[6] places greater emphasis on context effects. The theory suggests that a choice involving risk is strongly influenced by other contextual information. Unlike expected utility theory, prospect theory suggests that individuals do not take a comprehensive view of all possible outcomes that might be associated with a particular choice. Instead, they view the consequences of a particular choice as a change, which is judged as a gain or loss rather than the summation or final state of all possible outcomes. According to this theory, high probabilities of loss, such as certain death, might overcome low probabilities of the situation developing. Thus, prospect theory might suggest that fear of HIV should exceed fear of hepatitis B.

Methods

Subjects

The subjects were 264 undergraduate students at the University of California, San Diego, and San Diego State University. The students were enrolled in lower division courses in psychology and health sciences. These courses had not, at the time of the study, considered the AIDS epidemic. Sixty-one percent of the subjects were female and 39% were male. The mean age of the participants was 20.30 years (SD = 3.87 years).

TABLE 1—Comparison of Responses across Three Scenarios

Scenario	Total No. of Respondents	No. of Respondents Avoiding First Option	%
Disease (chance of infection following exposure × chance of death following infection)			
HIV (1% × 100%) vs hepatitis B (25% × 5%)	83	56	67
Insect (chance of being stung × chance of death after being stung)			
(1% × 100%) vs (25% × 5%)	91	58	64
(100% × 1%) vs (5% × 25%)	90	28	31

Note. HIV = human immunodeficiency virus. The disease scenario describes hypothetical HIV vs hepatitis B exposure. The insect scenarios compare hypothetical insect exposures.

Questionnaire

Each subject responded to one of three versions of a questionnaire. The first version described a situation explicitly specifying hypothetical hepatitis B and HIV exposures. Group 1 received a questionnaire that stated:

> Two diseases which can be acquired by accidental needlestick exposure to an infected patient are hepatitis B and AIDS. The risks of contracting hepatitis B from a needlestick are approximately 25%, and if one gets the disease there is approximately a 5% chance of dying. The risks of contracting AIDS from a needlestick exposure to an infected patient are approximately 1%, but 100% of those infected will die. Assume that a preventive vaccine is available for both diseases. You are a health care worker in a foreign country in which the prevalence of both diseases appears to be equal, but which can afford only one or the other vaccine for you. Which vaccine would you choose?

The subject could choose a hepatitis B or an AIDS vaccine.

The second and third groups were presented scenarios with similar contrasts in risks as in the first group, but in describing the hypothetical situations the scenarios eliminated the words *AIDS* and *hepatitis B*. The first of these conditions stated:

> You have been shipwrecked on the shore of a tropical country. The only passage to safety is by one of two routes, both of which expose you to insects of varying population intensity and lethality. Route A sends you to a valley where your likelihood of being stung is 25%. The chance of dying if you are stung is 5%. Route B sends you through a valley in which the insects are sparser but more lethal. Through this route you have a 1% chance of being stung, but if you are stung you are 100% certain to die. Which route would you choose?

The third set of scenarios (given to Group 3) was identical to the set given to the second group, except the probabilities of being stung and dying were reversed. Route A involved a 100% chance of being stung but a 1% chance of dying, while Route B involved a 5% chance of being stung and a 25% chance of dying.

Statistical Methods

Comparisons between conditions were achieved using χ^2 tests. *T* tests and χ^2 tests were used to evaluate the relationships between choice and demographic variables.

Results

Table 1 summarizes the choices for the three groups. Each choice was conceptualized as either avoiding HIV or avoiding hepatitis B risk situations. Because of the wording of the scenarios, this risk component was systematically varied across groups. In the insect scenario, 64% chose the route that avoided the 1% chance exposure and 100% death; in the explicit HIV–hepatitis B scenario, 67% preferred an AIDS over a hepatitis B vaccine. The difference between these two groups was nonsignificant ($\chi^2 = 0.26$).

When the probabilities in the insect scenarios were reversed (100% chance of being infected, 1% chance of dying; 5% chance of being infected, 25% chance of dying), the results changed. Under these circumstances, only 31% preferred the route with the 100% chance of being infected and 1% chance of dying. The difference between these two insect scenar-

ios (i.e., reversing the chances of being stung and dying) was statistically significant ($\chi^2 = 19.31, P < .001$). The subjects' response to the insect scenario with the reversed probabilities was also significantly different from the response to the explicit HIV–hepatitis B scenario ($\chi^2 = 22.85, P < .001$).

Age and sex did not significantly influence choice, as evaluated respectively by t tests (age) and χ^2 tests (sex). For example, in Group 2, 69% of the women and 66% of the men selected Choice 1. The mean age of those selecting Choice 1 was 21.61 years, while the mean age of those selecting Choice 2 was 20.34 years ($t_{89} = 1.33$, NS).

Discussion

Expected utility theory suggests that decisions should be based on the product of outcome and expectancy.[4,5] In this experiment we created scenarios in which the expected utilities of two outcomes were nearly equivalent. In all scenarios the subjects avoided outcomes that had a low probability of certain death. In the one condition that did not include certain death, subjects preferred the choice that offered the lowest probability of death. The effect of context observed in this study lends support to prospect theory.[6] In comparison to expected utility theory, prospect theory is better able to account for the effects of certain outcomes upon judgment.

Since responses to the insect scenarios and the identified disease scenario were not significantly different, data from this study may explain the differential attention to hepatitis B and HIV by health care providers. If fear of AIDS caused alarm independent of the probability information, we should have observed differential responses to the insect and identified disease scenarios with the same probabilities. Although the expected value of death due to exposure to HIV is about equivalent to the expected value of death from hepatitis B, the fear of HIV appears to exceed the fear of hepatitis B. Perhaps the most important explanatory variable is the probability of certain death. The chances of being infected with HIV are small, yet the consequences are very serious. Fear of living for any length of time infected with a virus for which the outcomes are always fatal apparently overrides the concern that the initial chances of being infected are small. Conversely, accidental exposure to hepatitis B, although highly infectious, offers the hope that even if infected the victim has only about a 5% chance of ultimately dying of the illness. Thus, what at first appears to be irrational behavior in the face of alternative risks might actually be rational. Who would not prefer to live, no matter how long or short a time, in a state of hopeful uncertainty when it comes to death (as most of us do every day) than to be forced to alter our perception and live in a state of dismal certainty?

The results might also explain fears associated with other low-probability events that are perceived to have certain mortal outcomes. For example, the fear of being involved in an airplane crash appears to greatly exceed the fear of becoming involved in an auto crash, even though the probability of injury and death from auto crashes exceeds that from airplane crashes, given that many more auto crashes occur. We hope these results also offer insights into public attitudes that might lead to practical approaches to education. For one thing, it seems apparent that the customary use of statistics in allaying anxieties and enforcing precautions must be improved by incorporating the "rationally irrational" fears experienced by most, if not all, human beings.

This study has a number of limitations. The scenarios were hypothetical, and participation was restricted to college students. It might be argued that fears of accidental exposure to HIV and hepatitis B are related to factors other than those that we studied. For example, there is the issue of medical secrecy. Although information about hepatitis B is typically listed in the patient record, information about the HIV status of patients is typically absent. Thus, providers can take precautions against hepatitis B but do not have the information to alert them to take precautions against HIV. We cannot discount this explanation; however, we point out that in this study the greater fears of HIV exposure over hepatitis B exposure were demonstrated by subjects who were not health care providers and were not familiar with patient records. Furthermore, the island scenario produced results almost identical to those of the HIV–hepatitis B scenario and did not mention the names of the diseases.

Another limitation of the study is that these scenarios are not entirely realistic. For example, the AIDS vs hepatitis B scenario asked which vaccine the respondent would prefer. Unfortunately, there is no real AIDS vaccine, making the generalization to the real-life situation difficult. However, we suggest that the apparently excessive fear of HIV exposure is not an irrational manifestation, but rather the rational reaction to the prospect of living with the certainty of proximate death. □

Acknowledgments

The authors thank Cathie Atkins, Bea Roppe, Liz Eakin, John Polich, and Amy Good for assistance in acquiring the data.

References

1. Centers for Disease Control. Guidelines for prevention of transmission of human immunodeficiency virus and hepatitis B virus to health-care and public-safety workers. *MMWR.* 1989;38(suppl 6):5–6.
2. Centers for Disease Control. Recommendations for protection against viral hepatitis. *MMWR.* 1985;34:313–324, 329–335.
3. Marcus R. CDC cooperative needlestick group: Surveillance of health care workers exposed to blood from patients infected with the human immunodeficiency virus. *N Engl J Med.* 1988;319:1118–1123.
4. Raiffa H, Schlaifer R. *Applied Statistical Decision Theory.* Cambridge, Mass: Harvard Business School; 1961.
5. Von Neumann J, Morgenstern O. *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton University Press; 1944.
6. Kahneman D, Tversky A. Choices, values and frames. *Am Psychol.* 1983;39:341–350.

What Is the Dentist's Occupational Risk of Becoming Infected with Hepatitis B or the Human Immunodeficiency Virus?

Eli I. Capilouto, DMD, ScD, Milton C. Weinstein, PhD, David Hemenway, PhD, and Deborah Cotton, MD, MPH

ABSTRACT

Surveys have shown that dentists are reluctant to treat persons infected with the human immunodeficiency virus (HIV). However, dentists are much more willing to treat patients with infectious hepatitis B virus (HBV). This study shows that the annual cumulative risk of infection from routine treatment of patients whose seropositivity is undisclosed is 57 times greater from HBV than from HIV, and that the risk of dying from HBV infection is 1.7 times greater than the risk of HIV infection, for which mortality is almost certain. (*Am J Public Health.* 1992;82:587–589)

Introduction

Practicing dentists are reluctant—and in many cases unwilling—to treat patients infected with the human immunodeficiency virus (HIV).[1,2] The frightening possibility of occupationally acquiring HIV may be the most significant factor in dentists' reluctance to treat potentially infected patients.[3,4] On the other hand, dentists are less concerned about becoming infected with infectious hepatitis, and are less frightened to treat patients with this disease.[2] Despite these differences in perceptions of risk and in concern over treatment of HIV patients versus infectious hepatitis patients, the risks associated with treatment of these patients have not been examined.

In this paper we quantify dentists' risk of occupationally acquiring HIV and hepatitis B virus (HBV) as a result of a percutaneous injury, the most likely and efficient means of occupational infection.[5–13] Specifically, we answer two questions: (1) What is the likelihood of infection from routine contact with asymptomatic patients, whose HIV or HBV status is likely to remain undisclosed in a routine medical history or dental exam? and (2) What is the risk of infection per known infected patient treated? The analysis addresses the risk of a general dentist who has neither received the hepatitis B vaccine nor acquired immunity from previous infection.

The answers to these questions take on greater importance following a report of a significant need for dental care among symptomatic HIV-infected patients[14] and disclosure of a single case in which five patients of an HIV-infected dentist were infected, presumably during the course of their dental treatment.[15]

Methods

Estimates of the dentist's risk of infection are made using the following deterministic model[16,17]:

$$\text{Cumulative Risk} = 1 - [p(1 - ib)^n + (1 - p)]^m,$$

in which the parameters that affect risk are represented as follows:

- p = the probability that any individual patient will be infected,
- i = the probability that the dentist will sustain a percutaneous injury while delivering therapy to an individual patient,
- b = the probability of viral transmission given a percutaneous injury,
- n = the number of visits by each patient, and
- m = the number of different patients treated.

In addition to the assumption of a random draw of patients, the model assumes the independence of all of the model's events. (See reference 17 for a complete discussion of the model.)

Parameter Estimates

Values for the parameter estimates in the model are derived from a variety of secondary data sources. The parameters and their corresponding estimates and references[1,18–34] are listed in Table 1. Mid-range estimates are used in baseline model fitting.

For the first research question (i.e., What is the infection risk from asymptom-

Eli I. Capilouto is with the University of Alabama School of Public Health, Lister Hill Center for Health Policy, Birmingham, Ala. Milton C. Weinstein, David Hemenway, and Deborah Cotton are with the Harvard School of Public Health, Boston, Mass; Deborah Cotton is also with the Department of Medicine, Beth Israel Hospital, Boston, Mass.

Requests for reprints should be sent to Eli Capilouto, DMD, ScD, School of Public Health, University of Alabama at Birmingham, University Station, Birmingham, AL 35294.

This paper was submitted to the Journal July 31, 1991, and accepted with revisions December 9, 1991.

TABLE 1—Parameter Estimates and Cumulative Annual Risk of Infection for Human Immunodeficiency Virus (HIV) and Hepatitis B Virus (HBV)

Parameter (symbol in model)	Estimate Used in Model	Range of Estimates	Reference(s)
Probability of Patient Infection (p)			
HIV (%)	0.28[a]	(0.16, 1.3)[a]	18–23
HBV (%)	0.4	(0.75)[b]	24, 25
Probability of Infection Given Injury (b)			
HIV (%)	0.3	(0.13, 0.70)[c]	26
HBV (%)	12.0[d]	(12.0, 22.0)[d]	27–32
Probability of Percutaneous Injury per Visit (i)	0.26[e]	(0.09, 0.43)[e]	1, 33, 34
Annual Number of Patients (m)[f]	1380		34
Annual Number of Visits per Patient (n)[f]	2		34
Cumulative Annual Risk			
HIV (%)	0.006		
HBV (%)	0.34		

[a]Model estimates of the number of HIV-infected persons in the United States vary from 400 000 to 1 million.[18] These two extreme estimates and a mid-range estimate (700 000 persons) were applied to the 1990 US Census population figure of 249 million persons to produce seroprevalence estimates of 0.16%, 0.28%, and 0.4%. Direct seroprevalence studies of newborns in New York City yielded the upper prevalence estimate of 1.3%.[22]

[b]Only two representative studies of HBV prevalence were found.

[c]These estimates represent the 95% confidence bounds of an analysis[24] that pooled results from numerous studies.

[d]These estimates are derived from two studies that tested the efficacy of hepatitis B immune globulin and immune serum globulin (ISG) in health careworkers exposed to hepatitis B virus via needle stick. Subsequent to these trials, the availability of more sensitive radio immune assay tests and the reanalysis of the sera collected in the trials led the researchers to conclude that ISG was essentially a placebo. Hence, infection likelihoods are derived from the groups that received ISG. (See references 31 and 32 for a complete explanation).

[e]The annual number of percutaneous injuries reported by dentists in three different retrospective studies were applied to the mean number of annual visits for dentists in general practice.

[f]Although ranges are available for these estimates, they are not important to the central purpose of this research, which is to examine the comparative likelihoods of infection from HIV and HBV.

TABLE 2—Cumulative Risk (in percent) for Human Immunodeficiency Virus (HIV) and Hepatitis B Virus (HBV) Infection According to Number of Known Infected Patients Treated

No. Infected Patients Treated	Cumulative Risk (%)	
	HIV	HBV
1	0.0007	0.03
10	0.008	0.3
100	0.08	3.1
1000	0.78	26.8
2000	1.54	46.4
5000	3.82	78.9

Note. The parameter probability estimates (in percent) used in model fitting were as follows: for HIV, i = 0.26, b = 0.3, p = 1, n = 1; for HBV, i = 0.26, b = 12.0, p = 1, n = 1.

atic patients whose serostatus is unknown?), parameter estimates on patient visits are based on annual figures so that an annual cumulative risk of infection is obtained. For the second research question (i.e., What is the infection risk per known infected patient treated?), the probability of patient infection (p) equals 1 in the model. The cumulative likelihood of infection is estimated for patient visits ranging in number from 1 to 5000.

Results of Model Fitting

In the case in which the susceptible dentist presumably is unaware of his patients' serostatus, the cumulative annual risk of infection is 57 times as great for HBV as for HIV (0.34% versus 0.006%; Table 1), a difference attributable to the higher prevalence of HBV in the patient population and the more efficient transmission rate.

Conducting a one-way sensitivity analysis of the model fitting for HBV risk sheds light on the model's validity. Substituting the higher 0.22 transmission rate raises the estimated annual risk of HBV infection from 0.34% to 0.63%. Interestingly, this risk rate is similar to reported annual incidence rates of 5.6 and 6.2 per 1000 person-years obtained in epidemiologic studies[35,36] of dentists before the availability of hepatitis vaccine.

Table 2 presents the cumulative risks of infection according to number of patient visits by infected patients. After 5000 visits, the likelihood of infection is 79% for HBV and 3.82% for HIV, as a result of HBV's higher transmission rates. The numbers in Table 2 also demonstrate the relative risks for dentists under a policy that would shunt patients infected with HIV to certain dentists. For example, a dentist who treats 2000 HIV-infected patients in a year would have a risk of infection 190 times greater than that of a dentist who treats only 10 such patients (1.54 percent versus 0.008 percent).

Discussion

The results of this analysis raise an interesting issue concerning dentists' perception of the relative occupational risks from these two infections. If a dentist exposed to HIV seroconverts, the likelihood of the dentist's developing acquired immunodeficiency syndrome (AIDS) is believed to increase with duration of infection. In fact, development of AIDS within 12 to 15 years appears almost certain, according to current model estimations.[37,38] However, this analysis suggests that in this occupational setting the mortality risk from HBV may be higher than the risk from HIV, and in many of the cases death will be more immediate. According to the Centers for Disease Control, 1.4% of all reported hepatitis B cases in 1985 ended in death within the year.[39] Among those 40 and older, the death rate was 3.1%. If, as expected, another 10% became chronic carriers, then their risk of death from hepatocellular carcinoma over the next 30 years would be approximately 15%.[40] Factoring the likelihood of acute and long-term mortality into the risk of primary infection yields a mortality risk of 0.01% from primary HBV infection, a risk 1.7 times greater than the risk of becoming infected with HIV (0.006%).

Apparently, many dentists do not recognize these relative risks. In a survey of dentists,[2] only 31% responded that they would be willing to treat HIV-infected patients, whereas 73% said they would treat hepatitis patients. The protection afforded by hepatitis immunization (40% of the respondents had been immunized) may partially explain the greater willingness to treat hepatitis patients. If so, given our findings of high mortality risk from HBV exposure, serious policy alternatives to increase the number of dentists immunized against hepatitis should be implemented. Another likely explanation underlying the greater willingness to treat hepatitis patients may

be the differences in preferences for the health states associated with HIV and HBV infection.[41] In other words, however dreadful, the stigma and morbidity of the health states of HBV disease are preferable to similar dimensions of HIV disease.

Finally, this analysis raises implications for the policy favored by most practicing dentists of referring HIV-infected patients to public health clinics, dental schools, and hospitals for treatment. According to the results shown in Table 2, such a policy would result in a maldistribution of the risk. Are dentists at high risk cognizant of their level of risk? Should dentists at high occupational risk be paid a wage premium for their incurred risk? If so, how much? The model used in this analysis can be helpful in providing answers to such questions. In addition, the model can be expanded to assess the reductions in risk that can be obtained from infection control proposals, and to weigh such benefits against their respective costs. □

Acknowledgments

The contributions of John Graham, PHD of the Harvard School of Public Health to earlier versions of this paper are gratefully acknowledged. Dr. Capilouto's training and contribution to this study was supported by a National Research Service Award, PHS Grant HS-000-20-03.

References

1. Verrusio AC, Neidle EA, Nash KD, Silverman S, Horowitz AM, Wagner KS. The dentist and infectious diseases: a national survey of attitudes and behavior. *J Am Dent Assoc.* 1989;118:553–562.
2. Sadowsky D, Kunzell C. Are you willing to treat AIDS patients? *J Am Dent Assoc.* 1990;122:29–32.
3. Gerbert B. AIDS and infection control in dental practice: changing dentists' knowledge, attitudes, and behaviors. *J Am Dent Assoc.* 1987;114:311–314.
4. Gerbert B, Badner V, Maguire B. AIDS and dental practice. *J Public Health Dent.* 1988;48:68–73.
5. Centers for Disease Control. Update: universal precautions for prevention of transmission of human immunodeficiency virus, hepatitis B virus, and other bloodborne pathogens in health-care settings. *MMWR.* 1988;37:377–388.
6. Villarejos VM, Visoná KA, Guitiérrez A, Rodriguez A, et al. Role of saliva, urine and feces in the transmission of type B hepatitis. *N Engl J Med.* 1974;291:1375–1380.
7. Groopman JE, Salahuddin SZ, Sarngadharan MG, et al. HTLV-III in saliva of people with AIDS-related complex and healthy homosexual men at risk for AIDS. *Science.* 1984;226:447–449.
8. Ho DD, Byington RE, Schooley RT, et al. Infrequency of isolation of HTLV-III virus from saliva in AIDS. *N Eng J Med.* 1985;313:1606.
9. Bancroft WH, Snitbhan R, Scott RM, et al. Transmission of hepatitis B virus to gibbons by exposure to human saliva containing hepatitis B surface antigen. *J Infect Dis.* 1977;135:79–85.
10. Scott RM, Snitbhan R, Bancroft WH, Alter HJ, Tingpalapong M. Experimental transmission of hepatitis B by semen and saliva. *J Infect Dis.* 1980;142:67–71.
11. Alter HJ, Purcell RH, Gerin JL, et al. Transmission of hepatitis B surface antigen-positive saliva and semen. *Infect Immun.* 1977;16:928–933.
12. Glasser JB, Nadler JP. Hepatitis B virus in a cardiopulmonary resuscitation training course: risk of transmission from a surface antigen-positive participant. *Arch Intern Med.* 1985;145:1653–1655.
13. Osterholm MT, Bravo ER, Crosson JT, et al. Lack of transmission of viral hepatitis type B after oral exposure to Hbsag-positive saliva. *Br Med J.* 1979;2:1263–1264.
14. Capilouto EI, Piette J, White BA, Fleishman J. Perceived need for dental care among persons living with AIDS. *Med Care.* 1991;28:745–754.
15. Centers for Disease Control. Update: transmission of HIV infection during invasive dental procedures—Florida. *MMWR.* 1991;40:377–381.
16. Weinstein MC, Graham JD, Siegel JE, Fineberg HV. Cost-effectiveness analysis of AIDS prevention programs: concepts, complications and illustrations. In: Turner CF, Miller HG, Moses LE, eds. *AIDS: Sexual Behavior and Intravenous Drug Use.* Washington, DC: National Academy Press; 1989:471–499.
17. Capilouto EI, Weinstein MC, Orav EJ, Hemenway D. Modeling dental health care workers' risk of occupational infection from bloodborne pathogens. *J Dent Educ.* 1990;54:687–688.
18. Hay JW. Projecting the medical costs of HIV/AIDS: an update with focus on epidemiology. In: *New Perspectives on HIV-Related Illnesses: Progress in Health Services Research.* Washington, DC: US Dept of Health and Human Services; 1989:84–97.
19. Centers for Disease Control. Human immunodeficiency virus infection in the United States: a review of current knowledge. *MMWR.* 1987;36(suppl S-6):1–48.
20. Centers for Disease Control. Prevalence of human immunodeficiency virus antibody in US active-duty military personnel, April 1988. *MMWR.* 1988;37:461–463.
21. Hoff R, Berardi VP, Weiblen BJ, Mahoney-Trout L, Mitchell ML, Grady GF. Seroprevalence of human immunodeficiency virus among childbearing women: estimating by testing samples of blood from newborns. *N Engl J Med.* 1988;318:525–530.
22. Novick LF, Berns D, Stricof R, Stevens R, Pass K, Wethers J. HIV seroprevalence in newborns in New York state. *JAMA.* 1989;261:1745–1750.
23. Karon JM, Dondero TJ, Curran JW. The projected incidence of AIDS and estimated prevalence of HIV infection in the United States. *J AIDS.* 1988;1:542–550.
24. Maynard JE. Nosocomial viral hepatitis. *Am J Med.* 1981;70:439–443.
25. McQuillan GM. *The Seroepidemiology of Hepatitis B Virus Infection in the United States, 1976–1980.* Baltimore, Md: The School of Hygiene and Public Health of The Johns Hopkins University; 1987. Doctoral thesis.
26. Henderson DK, Fahey BJ, Willy M, et al. Risk for occupational transmission of human immunodeficiency virus type 1 (HIV-1) associated with clinical exposures. *Ann Intern Med.* 1990;113:740–746.
27. Seeff LB, Wright EC, Zimmerman HJ, et al. Type B hepatitis after needle-stick exposure: prevention with hepatitis B immune globulin. Final report of the Veterans Administration Cooperative Study. *Ann Intern Med.* 1978;88:285–293.
28. Grady GF, Lee VA, Prince AM, et al. Hepatitis B immune globulin for accidental exposures among medical personnel: final report of a multicenter controlled trial. *J Infect Dis.* 1978;138:625–638.
29. Hoofnagle JH, Gerety RJ, Barker LF. Antibody to the hepatitis B surface antigen in immune serum globulin. *Transfusion.* 1975;15:408–413.
30. Grady GF, Rodman M, Larsen LH. Hepatitis B antibody in conventional gammaglobulin. *J Infect Dis.* 1975;132:474–477.
31. Hoofnagle JH, Seeff LB, Bales ZB, et al. Passive-active immunity from hepatitis B immune globulin. Reanalysis of a Veterans Administration cooperative study of needle-stick hepatitis. *Ann Intern Med.* 1979;91:813–818.
32. Grady GF, Cyr KM, Werner BG. Serologic reanalysis of the NHLBI needle-stick study. In: Szmuness W, Alter HJ, Maynard JE. *Viral Hepatitis.* Philadelphia, Pa: Franklin Institute Press; 1982:753–754.
33. Klein RS, Phelan JA, Freeman K, et al. Low occupational risk of human immunodeficiency virus among dental professionals. *N Engl J Med.* 1988;318:86–90.
34. American Dental Association. *The 1988 Survey of Dental Practice—Patient Visits to the Dentist.* Chicago, Ill: The Bureau of Economic and Behavioral Research; 1988.
35. Mosely JW, White E. Viral hepatitis as an occupational hazard of dentists. *J Am Dent Assoc.* 1975;90:992–997.
36. Segal HE, Evan C, Irwin GR, et al. Hepatitis B antigen and antibody in the United States Army: two-year follow-up of health care personnel. *Milit Med.* 1979;144:792–795.
37. Lagakos SW, De Gruttola V. The conditional latency distribution of AIDS for persons infected by blood transfusion. *J AIDS.* 1989;2:84–87.
38. Lifson AR, Hessol NA, Rutherford GW, et al. The natural history of HIV infection in a cohort of homosexual and bisexual men: clinical manifestation, 1978–1989. Presented at the Fifth International Conference on AIDS; June 1989; Montreal, Quebec, Canada.
39. *Hepatitis Surveillance Report No. 51.* Atlanta, Ga: Centers for Disease Control; 1987.
40. Beasley RP. Hepatitis B virus—the major etiology of hepatocellular carcinoma. *Cancer.* 1988;61:1942–1956.
41. Owens DK, Nease RF. Morbidity and mortality in the analysis of risk: effect of incorporation of preferences for diminished health states. Presented to the Society for Medical Decision Making; November 11, 1990; Boston, Mass. Abstract.

Prevalence of HIV Antibodies in Transsexual and Female Prostitutes

Baruch Modan, MD, DrPH, Reuven Goldschmidt, MD, Ethan Rubinstein, MD, Ami Vonsover, PhD, Marion Zinn, Rachel Golan, BSc, Angela Chetrit, BSc, and Tamar Gottlieb-Stematzky, PhD

ABSTRACT

Human immunodeficiency virus (HIV) prevalence was studied in an unselected group of 216 female and transsexual prostitutes. Subjects were asked about age, biological sex, marital status, children, length of occupation, sexual practices, and drug abuse history. Blood was drawn on site.

All 128 females who did not admit to drug abuse were seronegative; 2 of the 52 females (3.8%) who admitted to intravenous drug abuse were seropositive. In contrast, 11.1% of the 36 male transsexuals (including 3 out of 32 non–drug abusers) were seropositive.

The results support the notion that vaginal transmission of HIV is less effective than anal transmission. (*Am J Public Health.* 1992;82:590–592)

Introduction

We report the initial results of a survey aimed at defining the prevalence of human immunodeficiency virus (HIV) seropositivity in a population of transsexual and female prostitutes working side by side. To the best of our knowledge no data on HIV infection among transsexuals is available.

Method

During a 36-month period, a multidisciplinary team manned a van that visited the major location of open prostitution in the Tel Aviv area, located on a northern beach of the city. To cover as wide a proportion of the prostitute population as possible, we visited the locale at different hours of the day approximately once every 6 weeks during the first year and approximately every 6 months subsequently. Although prostitution is not legal in Israel, the laws against it are not usually enforced. According to unofficial police estimates, the study covered at least 70% of Tel Aviv prostitutes known to the police.

During each visit a short demographic questionnaire was filled out by an interviewer. Informed consent was also obtained. The questions asked included questions about age, sex, marital status, number of children, years on the job, intravenous drug abuse (IVDA), and sexual practices.

As reported by the study subjects during the interviews, the Tel Aviv transsexual prostitute community is considered to comprise between 35 and 40 persons who practice prostitution in the same locale. These are biologically born males who live and dress as women. Most of them have either undergone castration or are in the process of getting ready to have such an operation. They are indistinguishable physically from the female prostitutes and share the same client population. They all reported being on continuous estrogen maintenance. About 25% of those interviewed reported a sex-change operation that included the forming of an artificial vagina.

A sample of blood was drawn on site, and specimens were transferred to the Chaim Sheba Medical Center laboratories. All serum samples were tested blindly for antibodies to HIV by a commercial enzyme-linked immunosorbent assay (ELISA kit, ELAVIA, Institute Pasteur, France). All sera found to be positive by ELISA were confirmed by an immunoblot assay (Western blot technique) using a commercial kit (Dupont Biotech Research Laboratory, Ind., USA), interpreted according to the supplier's instructions.

Results

The 216 individuals examined included 180 biological women and 36 transsexuals who had been born males. Twenty-six percent of the females and 10.3% of the males reported previous or current IVDA. All of the biological males admitted to anal intercourse at some time during their sexual lives. In contrast, less than 10% of the female prostitutes admitted to having engaged in anal intercourse.

For the total group, ages ranged between 18 and 57, with a mean of approximately 30 years. There was no difference in mean age between females who admitted to drug abuse, the rest of the female group, and the transsexuals.

Baruch Modan, Reuven Goldschmidt, Marion Zinn, Rachel Golan, and Angela Chetrit are with the Department of Clinical Epidemiology; Ethan Rubinstein is with the Department of Infectious Diseases; and Ami Vonsover and Tamar Gottlieb-Stematzky are with the Department of Virology—all at Chaim Sheba Medical Center, Tel Hashomer, Tel Aviv University Medical School, Israel. Baruch Modan is currently with the National Center for Health Statistics, Centers for Disease Control, in Hyattsville, Md.

Requests for reprints should be sent to Baruch Modan, MD, DrPH, National Center for Health Statistics, 6525 Belcrest Road, Hyattsville, MD 20782.

This paper was submitted to the Journal May 13, 1991, and accepted with revisions November 12, 1991.

Fifty percent of the females and 42.9% of the males had practiced prostitution for 5 years or longer; 22.9% and 28.6%, respectively, had practiced it for more than 10 years. There was no difference in length of occupation between females who reported IVDA and those who did not. Condoms were used by both sexes during the study period. More detailed information on the frequency of this practice could not be obtained.

Table 1 presents the numbers and percentages of seropositive subjects by biological sex and reported history of drug abuse. Close to 11% of the transsexuals were found to be HIV positive, as compared with only 1.1% of the total group of females. Breakdown by drug abuse history shows that 3.8% of the females who admitted to IVDA were HIV positive. *All* females who denied a history of IVDA were seronegative. Among the transsexuals, there were 3 seropositive among the 32 who did not admit to IVDA. The differences between the transsexuals and the females are significant at the $P < .01$ level by Fisher's Exact Test.

To estimate the effect of IVDA and gender, data were stratified by these variables. The odds ratio for female gender was 20.8 (90% confidence interval, 4.4–101.7); the odds ratio for IVDA was 7.3 (90% confidence interval, 1.6–34.3).

TABLE 1—Age Distribution of the Study Population and Percentage Seropositive, by Sex and Drug Abuse

Category	Age				Seropositive	
	n	Mean	SD	Range	n	%
Females, no drugs	128	31.1	7.5	19–54	—	0.0
Females, drug history	52	28.1	5.9	19–40	2	3.8
Total females	180	30.2	7.2	19–54	2	1.1
Males	36	31.0	9.6	18–57	4	11.1

Note. Age was unclear in 1 female and 3 males.

Discussion

HIV has been isolated from genital secretions as well as from genital ulcers, saliva, semen, and seminal plasma of AIDS patients.[1–8] Yet, although the male to female route has never been disputed,[9–12] the exact mechanism of transfer has not been fully established.

The discrepancy between male and female prostitutes working side by side suggests that female prostitutes who are not drug abusers may be at less risk for HIV transmission than are transsexuals. There are a number of possible reasons for this discrepancy: a lower frequency of condom use, a possible increase in efficiency of transmission of HIV by an artificial vagina in the transsexual population (e.g., more friable mucosa), and a possible role of estrogen replacement. Still, most of the subjects in both the male and female groups reported full use of condoms at the time of the study.

The fact that all of the seropositive women admitted to drug abuse is in line with other studies of prostitutes in Western societies. A thorough search of recent literature fails to demonstrate unequivocal seropositivity among British,[13] French,[14] German,[15] Italian,[16] or Dutch[17] prostitutes without drug histories. In Greece,[18] 6% of 200 prostitutes examined were reported to be seropositive; however, their lack of drug histories may be questioned. In the same context, 5% seropositivity was observed among female prostitutes seen in selected venereal disease clinics in the United States.[19–21] Yet anal intercourse, genital ulcers,[21] or direct vaginal seminal contact[22] were not ruled out.

The data are also compatible with the rarity of female-to-male transfer of HIV,[22–28] and they may raise the question of whether female infection occurs primarily outside the vaginal route. Even if some of the prostitutes had falsely denied drug abuse, the resulting misclassification would have led to a bias in the opposite direction. Unfortunately, no male partners of these female prostitutes could be included in the study. To address the question of their infection status, one would have to prospectively follow a group of male partners sexually exposed to HIV-infected prostitutes. This mission would have been impossible to accomplish in a population that volunteered to participate in a study that promised anonymity. Because the transsexuals had more anal contacts in their pasts than did the female prostitutes, our findings may suggest that vaginal transmission of HIV is less effective than anal transmission. On the other hand, the role of more effective barrier precautions taken by some of the female prostitutes cannot be ruled out. □

Acknowledgment

We are highly indebted to Dr Shlomo Segev, Mrs Ayala Lusky, and Mr Shalom Frish for their invaluable aid.

References

1. Friedland GH, Klein RS. Transmission of the human immunodeficiency virus. *N Engl J Med.* 1987;317:1125–1135.
2. Archibald DW, Witt DL, Craven DE, et al. Antibodies to human immunodeficiency virus in cervical secretions from women at risk for AIDS. *J Infect Dis.* 1987;156:240–241.
3. Archibald DW, Zon L, Groopman JE, et al. Antibodies to human T-lymphotropic virus type III (HTLV-III) in saliva of acquired immunodeficiency syndrome (AIDS) patients and in persons at risk for AIDS. *Blood.* 1986;67:831–834.
4. Vogt MW, Witt DJ, Craven DE, et al. Isolation patterns of the human immunodeficiency virus from cervical secretion during the menstrual cycle of women at risk for the acquired immunodeficiency syndrome. *Ann Int Med.* 1987;106:380–382.
5. Kreiss JK, Coombs R, Plummer F, et al. Isolation of human immunodeficiency virus from genital ulcers in Nairobi prostitutes. *J Infect Dis.* 1989;160:380–384.
6. Zagouri D, Bernard J, Leibowitz J, et al. HTLV III in cells cultured from semen of two patients with AIDS. *Science.* 1984;226:449–451.
7. Ho D, Byington RE, Schooley RT, et al. Infrequency of isolation of HTLV-III virus from saliva in AIDS. *N Engl J Med.* 1985;313:1606.
8. Guinan ME, Hardy A. Epidemiology of AIDS in women in the United States—1981 through 1986. *JAMA.* 1987;257:2039–2042.
9. Stewart GJ, Tyoer JPP, Cunninghan AL, et al. Transmission of human T-cell lymphotropic virus type III (HTLV-III) by artificial insemination by donor. *Lancet.* 1985;2:581–585.
10. Pitchenik AE, Shafron RD, Glasser RM, Spira TJ. The acquired immunodeficiency syndrome in the wife of a hemophiliac. *Ann Int Med.* 1984;100:62–65.
11. Fischl MA, Dickinson GM, Scott GB, et al. Evaluation of heterosexual partners, children and household contacts of adults with AIDS. *JAMA.* 1987;257:640–644.
12. Piot P, Plummer EA, Mhala FS, Lamboray JL, Chin J, Mann JM. AIDS: An international perspective. *Science.* 1988;239:573–579.
13. Barton SE, Underhill GS, Gilchrist C, et al. HTLV-III antibody in prostitutes. *Lancet.* 1985;2:1424.

14. Brenky-Faudeux D, Fribourg-Blanc A. HTLV-III antibody in prostitutes. *Lancet.* 1985;2:1424.
15. Smith GL, Smith KF. Lack of HIV infection and condom use in licensed prostitutes. *Lancet.* 1986;2:1392.
16. Tirelli U, Vaccher E, Carbone A, et al. HTLV-III antibody in prostitutes. *Lancet.* 1985;2:1424.
17. Coutinho RA, van der Helm TH. Geen aanwijzing voor LAV/HTLV-III onder prostituees in Amsterdam die geen drugs gebruiken. *Ned Tijdschr Geneeskd.* 1986;130, no. II.
18. Papervangelou G, Roumeliotou-Karaysonis A, Kallinkos G, Papoutakis G. LAV/HTLV-III infection in female prostitutes. *Lancet.* 1985;2:1018.
19. Antibody to human immunodeficiency virus in female prostitutes. *JAMA.* 1987;257:2011–2013.
20. Redfield RR, Markham PD, Salahuddin SZ, et al. Heterosexually acquired HTLV-III/LAV disease (AIDS-related complex and AIDS). *JAMA.* 1985;254:2094–2096.
21. Marmor M, Weiss LR, Lyden M, et al. Possible female-to-female transmission of human immunodeficiency virus. *Ann Int Med.* 1986;105:969.
22. Wykoff RF. Female-to-male transmission of HTLV-III. *JAMA.* 1986;255:1704–1705.
23. Stoneburner RL, Chiasson MA, Weisfuse IB, Thomas PA. The epidemic of AIDS and HIV-1 infection among heterosexuals in New York City. *AIDS.* 1990;4:99–106.
24. Schultz S, Milberg JA, Kristal AR, Stoneburner RL. Female-to-male transmission of HTLV-III. *JAMA.* 1986;255:1703–1704.
25. Polk FB. Female-to-male transmission of AIDS. *JAMA.* 1985;254:3177–3178.
26. Ross MW. Prevalence of risk factors for HIV infection in the Australian population. *Med J Australia.* 1988;149:362–365.
27. Cameron DW, Simonsen JN, D'Costu LJ, et al. Female to male transmission of human immunoefficiency virus: risk factors for seroconversion in men. *Lancet.* 1989;2:403–407.
28. Padian NS, Shiboski SC, Jewell NP. Female to male transmission of human immunodeficiency virus. *JAMA.* 1991;266:1664–1667.

Predictors of Condom Accessibility among Hispanics in San Francisco

Barbara V. Marin, PhD, and Gerardo Marin, PhD

ABSTRACT

Ready access to condoms can be an important means of slowing the spread of sexually transmitted diseases and human immunodeficiency virus. We identified the factors associated with keeping or carrying condoms in a random-digit dialing survey of 522 Hispanic adults aged 18 to 65 years in San Francisco. Overall, more Hispanic men reported having condoms than Hispanic women (55.2% vs 23.8%). Hispanic men also reported much higher levels of risky sexual behavior than Hispanic women. Condom promotion efforts with Hispanics will require different approaches for each gender. (*Am J Public Health.* 1992;82:592–595)

Introduction

Sexually transmitted diseases (STDs) and acquired immunodeficiency syndrome (AIDS) are major health problems for Hispanics in the United States. While Hispanics make up approximately 8.5% of the population of the United States, they account for over 15% of AIDS cases,[1] and the incidence of AIDS attributable to heterosexual contact is 4 times higher for Hispanic men and 11 times higher for Hispanic women than for non-Hispanic Whites.[2] Currently, the rates of gonorrhea, syphilis, and chancroid are all many times higher for Hispanics than for non-Hispanic Whites in the United States.[3–5]

Appropriate condom use is an important method of preventing the spread of human immunodeficiency virus (HIV) and other STDs among Hispanics. Rates of condom use among married women, including Hispanics, remain low,[6] and rates of use for unmarried Hispanics are rarely reported. To explore and describe the demographic, attitudinal, and behavioral factors associated with condom use among Hispanics and to identify areas that should be emphasized in health education campaigns to promote condom use in this group, we surveyed a random sample of Hispanic adults in San Francisco by telephone.

Methods

Procedure

We used a modified Mitofsky-Waksberg method of random-digit dialing[7] to survey Hispanic adults in San Francisco 18 to 65 years of age. The refusal rate was 8.4%.

Measures

To assess condom accessibility, respondents were asked: "Do you keep condoms somewhere or carry them with you?" Also, they reported their frequency of condom use in the previous 12 months (always, often, sometimes, seldom, or never) and frequency of use with their primary sexual partner, with all male partners, and (in the case of men) with all female partners. Predictor variables included embarrassment (a lot or a little vs none) in buying, using, or, in the case of

Barbara V. Marin is with the Division of Clinical Epidemiology, Department of Epidemiology and Biostatistics, and Center for AIDS Prevention Studies, Institute for Health Policy Studies, University of California, San Francisco. Gerardo Marin is with the Department of Psychology, University of San Francisco.

Requests for reprints should be sent to Barbara V. Marin, PhD, Center for AIDS Prevention Studies, Box 0886, University of California, San Francisco, CA 94143-0886.

This paper was submitted to the Journal October 19, 1990, and accepted with revisions August 22, 1991.

women, asking a partner to use condoms; proportion of correct responses to 11 items about AIDS and HIV from the National Center for Health Statistics AIDS supplement to the National Health Interview Survey; and male respondents' beliefs about how commonly condoms are used by men in San Francisco. High-risk sexual behaviors in the 12 months prior to the survey were also predictors, defined as follows: "High-risk" men reported sexual relations with more than one female partner, with a prostitute, or with another man. "High-risk" women reported sexual relations with more than one male partner. Demographic predictors included employment status, income (yearly income of less than $15 601 vs $15 601 or more), age (less than 35 vs 35 or more), religiosity, and marital status (married or living together vs other). A five-item acculturation scale measured use of English and/or Spanish,[8] and the mean of responses was used to define low (1 through 2.99) and high (3.0 through 5.0) acculturation.

Analysis

Logistic regressions were conducted separately for each gender using the condom accessibility item as outcome and the other measures as predictors. To avoid problems with multicollinearity, education was not entered, as it correlated highly (.48) with acculturation.

Results

Five hundred twenty-two Hispanic adults in San Francisco were interviewed by telephone in autumn 1989. Table 1 includes relevant demographic statistics for the sample as a whole as well as for each gender. Table 2 presents specific sexual behaviors reported by the sample, and Table 3 indicates condom use among respondents reporting low- and high-risk sexual behaviors. Tables 4 and 5 present the results of the multiple logistic regressions separately for men and women. Table 6 includes information on where respondents obtain and keep condoms.

TABLE 1—Demographic Characteristics of the Total Sample and Each Gender Group

	Total Sample (n = 522)	Men (n = 220)	Women (n = 302)
Age, mean (SD)	35.6	34.4 (11.8)	36.4 (12.6)
Education, mean y (SD)	10.9	11.5 (4.5)	10.5* (4.4)
Employed last 2 weeks, %	63.0	77.7	52.3***
Family income more than $300 per week, %	57.2	52.7	60.5
Married or living with someone, %	61.0	55.6	65.0*
Less acculturated, %	73.3	69.5	76.1
Ancestry, %			
Mexican	33.7	36.8	31.6
Central American	53.6	51.4	55.1
Other	12.7	11.8	13.3

Note. Significance tests were for gender differences.
*$P < .05$.
**$P < .01$.
***$P < .001$.

TABLE 2—Specific Sexual Attitudes and Behaviors Reported

	Men (n = 220), %	Women (n = 286), %
Embarrassed to buy condoms	32.1	57.7
Sexual behaviors in prior 12 months		
Multiple heterosexual partners	20.3	4.6
Homosexual relations	3.6	...
Sex with a prostitute	5.7	...

TABLE 3—Condom Use by Risk Status for Hispanic Men and Women

	Men				Women			
	Low Risk		High Risk		Low Risk		High Risk	
	%	n	%	n	%	n	%	n
Always use with primary partner	12.5	128	28.0	25	12.9	206	9.1	11
Always use with heterosexual partners	...	...	50.0	48	...	...	21.4	14

ated with prostitution and extramarital relationships in Latin America.[9] However, only a third of men at risk reported "always" using condoms with all partners. Culturally appropriate campaigns to promote condoms to Hispanic men may be most successful if targeted to those with multiple partners.

Hispanic women, especially the less acculturated, often did not keep or carry condoms. A small group of women (5%) reported placing themselves at risk through multiple partners, and these women were somewhat, though not significantly, more likely to carry condoms. An unknown, but probably larger, number of Hispanic women may be placed at risk when their partners engage in risky behaviors. These women are likely to resist condom promotion efforts, since they reported that they were often embarrassed to buy or ask their partner to use condoms and that they were more likely to obtain condoms from health centers or friends than to purchase them.

Traditional Hispanic culture emphasizes the need for men to express their sexuality and for women to avoid such expression.[10,11] Both men and women often reported being embarrassed to buy condoms, especially those who reported not having condoms. Embarrassment has

Discussion

In this study, more Hispanic men reported having condoms available than Hispanic women. Hispanic men also reported much higher levels of risky sexual behavior than Hispanic women. Those men who had multiple sexual partners were more likely to have condoms available. This is not entirely surprising, since condom use has traditionally been associ-

TABLE 4—Multiple Logistic Regression Results: Percentage of Hispanic Men Keeping or Carrying Condoms

	n	%	Odds Ratio	95% Confidence Interval
Overall	220	55.2		
Sexual behaviors				
High risk	55	85.5	4.65	1.99–10.85
Low risk	157	44.6		
Age				
<35	130	66.9	0.36	0.18–0.71
≥35	90	37.8		
Embarrassed to buy condoms				
No	149	58.4	0.42	0.21–0.86
Yes	70	47.1		
Marital status				
Has regular partner	120	45.0	0.63	NS
No regular partner	96	66.7		
Acculturation				
More	67	61.2	1.28	NS
Less	153	52.3		

TABLE 5—Multiple Logistic Regression Results: Percentage of Hispanic Women Keeping or Carrying Condoms

	n	%	Odds Ratio	95% Confidence Interval
Overall	297	23.8		
Sexual behaviors				
High risk	14	42.9	1.81	NS
Low risk	276	22.8		
Age				
<35	157	34.4	0.22	0.11–0.44
≥35	139	10.8		
Embarrassed to ask partner to use condoms				
No	189	29.1	0.39	0.19–0.80
Yes	104	13.5		
Marital status				
Has regular partner	194	27.8	2.97	1.41–6.26
No regular partner	102	14.7		
Acculturation				
More	70	44.3	4.50	2.13–9.49
Less	226	16.8		

TABLE 6—Where Respondents Obtain and Keep Condoms

	Men (n = 121), %	Women (n = 69), %
Obtain		
Purchase	80.9	60.3
Health center or gift	19.1	39.7
Keep		
In wallet/purse	29.8	17.4
At home only	63.6	79.7
At work/ elsewhere	6.6	2.9

often been associated with inconsistent condom use in other populations,[12,13] suggesting that interventions to increase comfort with sexual topics could be useful adjuncts to condom promotion campaigns among Hispanics. Given the low proportion of less-acculturated (i.e., Spanish-speaking) women who had condoms, these women may be in special need of such an intervention.

This study has important limitations. Hispanics in San Francisco are different demographically from those in the Northeast, who tend to be predominantly from Puerto Rico and other Caribbean countries. The HIV/AIDS epidemic has affected a disproportionate number of Puerto Ricans in the northeastern United States,[14,15] although levels of STDs are high among Hispanics in most areas of the United States. An additional limitation of most studies of condom use is that self-reports of sexual behaviors can produce embarrassment and a desire to misrepresent actual behavior.

The high rates of sexual activity outside the primary relationship for Hispanic men in this study suggest that AIDS/STD prevention interventions targeting Hispanic men might include a component designed to improve sexual satisfaction within the primary relationship. In one study, Hispanic women, especially those who were married, reported low frequency of sexual intercourse and low sexual satisfaction.[16] For women with partners who engage in high-risk activities, interventions are needed to increase the awareness of their own risk and to improve their sexual communication skills; also, there is a need for the development, testing, and promotion of methods of protection that are not dependent on men's behavior.[17] □

Acknowledgments

This research was funded with Grants DA04928 from the National Institute on Drug Abuse and MH46777 from the National Institute of Mental Health.

An earlier version of this paper was presented at the Sixth International Conference on AIDS, June 1990, San Francisco, Calif.

The authors wish to acknowledge the thoughtful comments of Drs. Eliseo J. Pérez-Stable and Norman Hearst on an earlier draft of this paper and the contributions of Rolando Juarez and Samuel Posner to the data collection and analysis.

References

1. Centers for Disease Control. *HIV/AIDS Surveillance.* Atlanta, Ga: Centers for Disease Control; 1990.
2. Holmes KK, Karon JM, Kreiss J. The increasing frequency of heterosexually acquired AIDS in the United States, 1983–1988. *Am J Public Health.* 1990;80:858–862.
3. Aral S, Holmes KK. Epidemiology of sexually transmitted diseases. In: Holmes KK, Cates W Jr, Lemon SM, Stamm WE, eds. *Sexually Transmitted Diseases.* New York, NY: McGraw-Hill; 1990:126–141.
4. Centers for Disease Control. Syphilis and congenital syphilis—United States, 1985–1988. *MMWR.* 1988;37:486–489.
5. Schmid GP, Sanders LL, Blount JH, Alexander ER. Chancroid in the United States: reestablishment of an old disease. *JAMA.* 1987;258:3265–3268.
6. Mosher WD, Bachrach CA. *Contraceptive Use, United States, 1982.* Hyattsville, Md: National Center for Health Statistics. DHHS publication no. PHS 83-1988 (Vital and health statistics; series 23; no. 12).
7. Marin G, Marin B, Pérez-Stable EJ. Feasibility of a telephone survey to study a minority community: Hispanics in San Francisco. *Am J Public Health.* 1990; 80:323–326.
8. Marin G, Sabogal F, Marin B, et al. De-

velopment of a short acculturation scale for Hispanics. *Hispanic J Behavior Sci.* 1987;9:183–205.
9. Bailey J, Lopez-Escobar G, Estrada A. A Colombian view of the condom. *Stud Fam Plann.* 1973;4:60–64.
10. Pavich EG. A Chicano perspective on Mexican culture and sexuality. *J Soc Work Hum Sexuality.* 1986;4:47–65.
11. Vasquez-Nuttall E, Romero-Garcia I, De Leon B. Sex roles and perceptions of femininity and masculinity of Hispanic women. *Psychol Women Q.* 1987;11:409–425.
12. MacDonald NE, Well GA, Fisher WA, et al. High-risk STD/HIV behavior among college students. *JAMA.* 1990;263:3155–3159.
13. Stall R, Heurtin-Roberts S, McKusick L, Hoff C, Lang SW. Sexual risk for HIV infection among singles-bar patrons in San Francisco. *Med Anthropol Q.* 1990;4:115–128.
14. Bakeman R, McCray E, Lumb JR, Jackson RE, Whitley PN. The incidence of AIDS among Blacks and Hispanics. *J Natl Med Assoc.* 1987;79:921–928.
15. Selik RM, Castro KG, Pappaioanou M. Racial/ethnic differences in risk of AIDS in the United States. *Am J Public Health.* 1988;78:1539–1545.
16. Amaro H. Women in the Mexican-American community. *J Community Psychol.* 1988;16:6–20.
17. Stein ZA. HIV prevention: the need for methods women can use. *Am J Public Health.* 1990;80:460–462.

Legal Needle Buying in St. Louis

Wilson M. Compton III, Linda B. Cottler, Scott H. Decker, Douglas Mager, and Roosevelt Stringfellow

ABSTRACT

This study sought to determine if and why barriers to the over-the-counter purchase of syringes in the St. Louis metropolitan area might exist, given that no ordinance prohibits such a sale there. Two male research assistants (one African American, one White) approached 33 of the area's pharmacies to buy syringes. In 14 of those pharmacies, either the purchase was refused or the minimum number of syringes that could be bought was so large (at least 100) that the sale was not practical. Racial bias in rates of refusal and implications for prohibiting or restricting legal availability of syringes are discussed. (*Am J Public Health.* 1992;82:595–596)

Introduction

International studies have shown that the legal purchase of sterile syringes may be an important factor in reducing the spread of human immunodeficiency virus (HIV). On the other hand, obstacles to the purchase of sterile syringes may contribute to the rapid spread of HIV infection.[1,2] In France, the liberalization of regulations for purchasing syringes resulted in lower rates of needle sharing and increased use of sterile syringes by intravenous drug users (IVDUs).[3] In Edinburgh, Scotland, however, the rapid spread of HIV among IVDUs coincided with a police crackdown on the supply of legally available syringes.[4]

The State of Missouri, like 38 other states, has no ordinance prohibiting the over-the-counter sale of sterile syringes. Such legal and apparently easy availability of sterile syringes may be one of the reasons HIV infection rates are low (approximately 3%) among St. Louis' IVDU population[5] compared with rates in New York or New Jersey (50% to 60%).[6–8]

To describe the availability of legal needles, we designed a project to mimic the purchase of syringes by IVDUs at pharmacies throughout the St. Louis area, particularly in neighborhoods where intravenous drug use is prevalent. This project began at a time when we heard that local pharmacists had their own "War on Drugs" campaign, which included refusing to sell syringes to customers without proof of medical need.

Methods

A comprehensive list of pharmacies (n = 360) was drawn from the listings in the local Yellow Pages. To provide a widespread distribution, sample pharmacies were selected that were maximally distant from one another within defined geographic areas. A map of the St. Louis area was divided into 10 comparably sized areas, using major thoroughfares as dividing lines. One southwest suburban sector was excluded from sampling because police data indicated little drug-related activity. Four pharmacies were then selected within each sector, for a total of 36 pharmacies scheduled for sampling within nine geographical sectors.

During field work, several pharmacies were found to be out of business, so the nearest available pharmacy was substituted. Of the 36 selected pharmacies, 33 were visited during one week in May 1990 for a 92% completion rate. Of those visited, 25 were part of major chains; 8 were independent or affiliated with minor chains, hospitals, or medical centers.

The two male research assistants who carried out the buying project were 41 and 46 years old. One was African American and one was White. Their attire dur-

Wilson M. Compton, Linda B. Cottler, Douglas Mager, and Roosevelt Stringfellow are with the Department of Psychiatry at the Washington University School of Medicine in St. Louis, and Scott H. Decker is with the Administration of Justice at the University of Missouri-St. Louis.

Requests for reprints should be sent to Wilson M. Compton III, MD, Department of Psychiatry, Washington University School of Medicine, 4940 Audubon Avenue, St. Louis, MO 63110.

This paper was submitted to the Journal December 17, 1990, and accepted with revisions September 16, 1991.

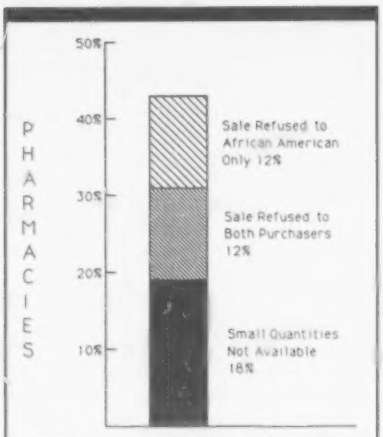


FIGURE 1—Obstacles to needle purchase.

ing the project was casual. They did not discuss the needle-buying project with pharmacy personnel and did not claim to be drug users. In addition, each was blinded as to the other's outcome.

Each research assistant approached the same pharmacy on consecutive days at approximately the same time of day and attempted to purchase syringes by saying "I need a pack of 28-gauge, 100-unit insulin syringes, please." A pack referred to a 10 pack. This size and type of syringe was chosen because a San Francisco study and our ethnographic consultant confirmed that it is the size used most often by IVDUs.[9]

The data collected included location and name of store; time of day and date when purchase was attempted; race, sex, and estimated age of pharmacy staff person confronted; and cost of syringes (if purchased). If the purchase was refused, the researchers engaged the pharmacist/clerk in casual conversation to determine the reasons for refusal.

Results

As shown in Figure 1, eight pharmacies (24%) refused to sell syringes; six others (18%) would not allow small quantities of syringes to be purchased. This meant that the purchase of a 10 pack or less was possible at 19 pharmacies (58%). Eight pharmacies refused to sell without confirmation of a medical condition; of those, four refused to sell syringes to both research assistants and four refused the African American only.

The predominant reason given for refusal to sell was simply "store policy." At certain stores, however, other reasons were given, which confirmed our original suspicions that pharmacists were waging their own War on Drugs. Two pharmacists reported that selling syringes is "too much trouble." Another reported that only a few specific customers could buy syringes and only by prior arrangement. A fourth reported that he would not sell syringes without a prescription because of the high incidence of drug abuse in the neighborhood. A fifth stated that he would not sell to "drug users."

The cost of syringes varied from $1.92 to $4.28 for a 10 pack. In six pharmacies, a minimum of 100 syringes could be purchased, at a cost of about $25; however, such a sale would be impractical for most drug addicts. In two pharmacies, single syringes could be purchased.

Conclusions

This needle-buying project provides baseline data concerning the availability of hypodermic syringes to the IVDU in a large midwestern city where no laws restrict the sale of such items. Almost half the pharmacies (42%) either refused to sell syringes or sold them only in costly quantities.

Evidence of possible racial bias in the sale of syringes to the white researcher but not to the African-American researcher was found and may be particularly important for African-American IVDUs. It is ironic that one of the populations most in need of prevention measures to combat the spread of acquired immunodeficiency syndrome (AIDS) is denied access to sterile syringes even when access is not legally denied.[10] African-American IVDUs may therefore be placed at higher risk for contracting AIDS because of racial bias in availability of sterile syringes. Users, especially if African American, may have little option but to sharpen, share, and reuse syringes. □

Acknowledgments

This study was funded in part by grant DA-06163 from the National Institute on Drug Abuse. However, the purchase of syringes was funded through private, nongovernmental funds.

References

1. Saxon AJ, Colsyn DA, Whittaker S, et al. Needle obtainment and cleaning habits of addicts. In: Harris LS, ed. *Problems of Drug Dependence 1989.* Washington, DC: US Govt Printing Office; 1990:418. National Institute on Drug Abuse Research Monograph 95. Department of Health and Human Services publication ADM 90-1663.
2. Ron A, Rogers DE. AIDS in New York City: the role of intravenous drug users. *Bull N Y Acad Med.* 1989;7:787–800.
3. Ingold FR, Ingold S. The effects of the liberalization of syringe sales on the behaviour of intravenous drug users in France. *Bull Narc.* 1989;41:67–81.
4. Robertson JR, Bucknall ABV, Welsby PD, et al. Epidemic of AIDS-related virus (HTLV-III/LAV) infection among intravenous drug users. *Br J Med.* 1986;292:527–529.
5. Cottler LB, Compton WM, Rice R, et al. Low rates of HIV infection among substance abusers reporting high-risk behavior: trend or time bomb? Presented at the 53rd Annual Scientific Meeting of the Committee on Problems of Drug Dependence; June 1991; Palm Beach, Fla.
6. Hopkins W. Needle sharing and street behavior in response to AIDS in New York City. In: Battjes RJ, Pickens RW, eds. *Needle Sharing among IVDA: National and International Perspectives.* Washington, DC: US Govt Printing Office; 1988:18–27. National Institute on Drug Abuse Research Monograph 80. Department of Health and Human Services publication ADM 88-1567.
7. Raymond CA. US cities struggle to implement needle exchanges despite apparent success in European cities. *JAMA.* 1988;260:2620–2621.
8. Hahn RA, Onorato IM, Jones TS, et al. Prevalence of HIV infection among intravenous drug users in the United States. *JAMA.* 1989;261:2677–2684.
9. Case P, Meredith G, Garcia D, et al. Needle exchange: from civil disobedience to public policy. *Multicultural Inquiries and Research into AIDS.* 1990;4(4):1–3.
10. D'Aquila RT, Peterson LR, Williams AB, Williams AE. Race/ethnicity as a risk factor for HIV-1 infection among Connecticut intravenous drug users. *J AIDS.* 1989;2:503–513.

Risk Factors Associated with Antibodies to Leptospires in Inner-city Residents of Baltimore: A Protective Role for Cats

James E. Childs, ScD, Brian S. Schwartz, MD, MS, Tom G. Ksiazek, PhD, DVM, R. Ross Graham, PhD, DVM, James W. LeDuc, PhD, and Gregory E. Glass, PhD

ABSTRACT

Leptospiral antibody prevalence was 16% in residents of Baltimore. Seropositivity was associated with age, gender, race, and bird ownership, and negatively associated with contact with rat excrement and cat ownership. Current cat ownership reduced the antibody risk associated with age and race from odds ratios (95% confidence interval) of 3.3 (2.0, 5.5) and 3.3 (1.1, 9.3), respectively, to the baseline level. These data establish the high prevalence of leptospiral antibody in Baltimore and suggest a protective role for cats in reducing the risk of human infection. (*Am J Public Health.* 1992;82:597–599)

Introduction

Leptospirosis is an acute, febrile spirochetal disease most commonly affecting the liver and kidneys. Although leptospirosis can be debilitating, infection frequently runs a milder or even asymptomatic course that may go undiagnosed.[1]

Leptospires (*Leptospira interrogans*) infecting humans are primarily maintained by warm-blooded vertebrates.[2] In cities, the Norway rat (*Rattus norvegicus*)[3–7] and dog[8] have been regarded as major reservoirs. The degree of human–animal contact with wild and domestic animals may be an important determinant of infection. In this paper we describe the prevalence of antibodies to leptospires in an inner-city population and examine risk factors associated with seropositivity.

Methods

Study Population

Sera and questionnaires were obtained from persons visiting a sexually transmitted disease clinic between March 1987 and December 1988, as described previously.[9,10] Subjects were primarily male (71.4%), Black (94.3%), and young (median age = 23.5 years); 99% lived within the city limits.[10]

Data Collection

Respondents signed informed consent forms and were individually interviewed. Verbal responses were recorded to questions of age; history of residence; yearly income; occupation; travel; frequency of rat and mouse sightings within current and past residences, on current or past streets and alleys, and at current or past work place; history of contact with rodents through trapping, excrement, handling animals, or receiving rodent bites; and current or past ownership or living with cats, dogs, or other pets (classified as mammal, bird, or amphibian/reptile/fish). Observed race and sex were noted. Blood from respondents was centrifuged and serum stored at −20°C.

Serology

Sera were screened for immunoglobulin G (IgG) by enzyme-linked immunosorbent assay (ELISA) using the serovar mwogolo, prepared by standard methods,[11] as antigen. Tests using this antigen are genus specific (R. R. Graham, unpublished data 1986). Sera reactive at 1:200 dilution were considered positive. A subsample of 37 sera positive by IgG ELISA were also tested by the microscopic agglutination test (MAT)[12] and screened for immunoglobulin M (IgM) by ELISA. Serogroups (serovars) used in MAT included Australis (*australis*), Autumnalis (*autumnalis*), Ballum (*ballum*), Bataviae (*bataviae*), Canicola (*canicola*), Grippotyphosa (*grippotyphosa*), Sejroe (*hardjo*), Icterohaemorrhagiae (*icterohaemorrhagiae, copenhageni*), Pomona (*pomona*), Pyrogenes (*pyrogenes*), Shermani (*shermani*), and Tarassovi (*tarassovi*). A MAT titer of 1:50 was considered positive for all serovars except *shermani*, where 1:800 was used because of the potential autoagglutination of the antigen.

James E. Childs and Gregory E. Glass are with the Department of Immunology and Infectious Diseases, and Brian S. Schwartz is with the Department of Environmental Health Sciences, The Johns Hopkins School of Hygiene and Public Health, Baltimore, Md. Tom G. Ksiazek and James W. LeDuc are with the Disease Assessment Division, United States Army Medical Research Institute of Infectious Diseases, Fort Detrick, Frederick, Md. R. Ross Graham is with the United States Army Medical Research Unit, Republic of Korea.

Requests for reprints should be sent to James E. Childs, ScD, Department of Immunology and Infectious Diseases, The Johns Hopkins University School of Hygiene and Public Health, 615 N. Wolfe Street, Baltimore, MD 21205.

This paper was submitted to the Journal April 5, 1991, and accepted with revisions July 18, 1991.

TABLE 1—Crude and Adjusted Odds Ratios for Variables Associated with Leptospiral Antibody in Baltimore, Md, 1987–1988

	Crude Odds Ratio (95% Confidence Interval)		Adjusted[a] Odds Ratio (95% Confidence Interval)	
Age (>19 vs ≤19)	2.3	(1.5, 3.5)	2.8	(1.8, 4.5)
Sex (male vs female)	1.5	(1.0, 1.9)	1.5	(1.0, 2.2)
Race (Black vs other)	1.8	(0.8, 3.9)	2.0	(0.9, 4.5)
Income (≤$8000 vs >$8000)	1.2	(0.8, 1.7)	1.6	(1.1, 2.4)
Contact with rat excrement (some vs none)	0.7	(0.5, 1.1)	0.5	(0.3, 0.9)
Current cat ownership (none vs ≥1)	2.1	(1.1, 3.8)	1.9	(1.1, 3.5)
Bird ownership (ever vs never)	1.4	(0.8, 2.5)	1.7	(1.0, 3.1)

[a]Adjusted for all other variables in table.

TABLE 2—Adjusted[a] Odds Ratios Associated with Age, Race, and Sex after Stratification by Current Cat Ownership

	Current Cat Ownership			
	None		One or More	
	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval
Age (>19 vs ≤19)	3.3	2.0, 5.5	0.8	0.3, 2.8
Race (Black vs other)	3.3	1.1, 9.3	0.4	0.1, 1.6
Sex (male vs female)	1.5	1.0, 2.3	1.4	0.4, 1.6

[a]Adjusted for age, sex, race, income, contact with rat excrement, current cat ownership, and bird ownership.

All ELISA positive and 163 negative sera were tested for cross reactivity to *Treponema pallidum* by a standard rapid plasma reagin (RPR) assay.[13]

Data Analysis

Analyses were performed with SAS®[14] or BMDP®[15] software systems. Variables were dichotomized to ease interpretation of results and to allow calculation of odds ratios. Crude odds ratios were calculated for the association between all study variables and leptospiral serological status. Stratification and logistic regression were used to assess interactions between variables and to control for potential confounding variables.[16] Variables thought to be important a priori were included in initial logistic regression models, and the contribution of other variables was assessed by partial F tests. After selection of a model, all possible interactions between variables were assessed and retained if statistically significant ($P < .05$) or if felt to be biologically important.

Results

The prevalence of ELISA antibodies to leptospires was 16% (185/1150), with a single positive IgM titer found. Titers by MAT ≥1:50 to one or more serovars were found in 31 of 37 ELISA positive sera. The most common serotypes eliciting a response were Batavia (20/37) and Icterohaemorrhagiae (16/37). There was no indication of significant cross-reactivity with *T. pallidum* ($\chi^2 = 1.88$, $P = .17$).

Crude (unadjusted) analyses showed associations of leptospiral antibody with age and male gender, but not with other demographic variables. Current ownership of cats was significantly associated with lower antibody prevalence. None of the rodent exposure variables were associated with antibody.

Logistic regression indicated that age (≥19), sex (male), Black race, and low income (≤$8000) were independent risk factors associated with leptospiral antibody (Table 1). Self-reported contact with rat excrement was negatively associated with seropositivity.

History of bird ownership (chickens or pet birds) was positively associated, and cat ownership negatively associated, with antibody. An interesting interaction was observed between cat ownership and both age and race: Age and race were risk factors for seropositivity only in persons who did not own cats (Table 2).

Discussion

This study documented a high prevalence (16%) of leptospiral antibody in an inner-city population. Associations between antibody and age and sex were expected based on studies of leptospirosis in the United States.[17] However, the data also indicated a potential protective role for cat ownership in reducing the risk of infection.

Studies conducted in Detroit have also reported high levels of leptospiral antibody (≥30%) in inner-city children under 6 years old.[18] As in Baltimore, with three cases of leptospirosis documented since 1987 (R. Dunning, written communication 1991), disease was uncommon in Detroit. Leptospirosis is a self-limited condition with usual spontaneous recovery within 4 to 9 days.[19] Afflicted persons may not seek medical care, and the disease also is underreported.[19]

The serological tests we used indicated few recent infections. Specificity of our ELISA assay was indicated by the 84% concordance with MAT and lack of cross-reactivity with RPR tests. Other serosurveys have found that up to 35% of ELISA positive sera will not react by MAT.[20]

Risk of leptospirosis associated with animals has rarely been explored beyond incriminating a species during a case investigation or epidemic,[21,22] or stratification of human antibody prevalence based on animal contact.[23,24] We found a lowered risk of leptospiral antibody for persons reporting past history of exposure to rat excrement. This study variable may have measured rat awareness rather than "risky" behavior. The clearest and most interesting association was between decreased leptospiral seropositivity and cat ownership. Cats could reduce the risk of rodent-borne leptospirosis if they reduced human contact with the reservoir. The relationship between cat ownership and race is also interesting, as Blacks in Baltimore own fewer cats than do Whites.[25] This difference in cat ownership could place Black residents at increased risk for contact with rodents and the zoonoses

they carry, and could contribute to the increased prevalence of antibody in Black residents.

Cats have been linked to human infection with several pathogens (e.g., *Toxoplasma gondii*[26] and *Yersinia pestis*[27]), but, to our knowledge, never in a protective role. These novel observations deserve further research and confirmation by other studies. □

Acknowledgments

This research was supported in part by contract DAMD17-89-C-9093 from the United States Army Medical Research and Development Command.

References

1. Peter G. Leptospirosis: a zoonosis of protean manifestations. *Pediatr Infect Dis.* 1982;1:282–288.
2. Torten M. Leptospirosis. In: Steele JH, ed. *Handbook Series in Zoonoses.* Boca Raton, Fla: CRC Press; 1979:363–421.
3. Thiermann AB. Incidence of leptospirosis in the Detroit rat population. *Am J Trop Med Hyg.* 1977;26:970–974.
4. Lewis M. The incidence of *Leptospira icterohemorrhagiae* in trapped rats in Philadelphia. *Am J Trop Med Hyg.* 1942;22:571–576.
5. Walch EW, Walch-Sordrager GB. Observations on *Leptospira icterohemorrhagiae* in wild rats of Baltimore. *Am J Hyg.* 1927;7:393–407.
6. Li H-Y, Davis DE. The prevalence of *Leptospira* and *Salmonella* in Norway rats of Baltimore. *Am J Hyg.* 1952;52:90–100.
7. Monlux WS. I. The incidence of leptospirosis in the rat population of Ithaca, New York, and vicinity. *Cornell Vet.* 1948;38:57–69.
8. Thiermann AB. Canine leptospirosis in Detroit. *Am J Vet Res.* 1980;41:1659–1661.
9. Childs JE, Glass GE, Ksiazek TG, Rossi CA, Barrera Oro JG, Leduc JW. Human-rodent contact and infection with lymphocytic choriomeningitis and Seoul viruses in an inner-city population. *Am J Trop Med Hyg.* 1991;44:117–121.
10. Childs JE, Glass GE, LeDuc JW. Rodent sightings and contacts in an inner-city population of Baltimore, Maryland, USA. *Bull Soc Vector Ecology.* In press.
11. Hartman EG, Van Houten M, Van der Donk J, Frik JF. Serodiagnosis of canine leptospirosis by solid-phase enzyme-linked immunosorbent assay. *Vet Immun Immunpathol.* 1984;7:33–42.
12. Cole JR, Sulzer CR, Pursell AR. Improved microtechnique for leptospiral microscopic agglutination test. *Appl Microbiol.* 1973;25:976–980.
13. National Communicable Disease Center. *Manual of Tests for Syphilis.* Atlanta, Ga: National Communicable Disease Center; 1969. DHEW publication number (PHS) 411.
14. SAS Institute Inc: *SAS User's Guide: Statistics. Version 5 Edition.* Cary, NC: SAS Institute Inc; 1985.
15. Dixon WJ, ed. *BMDP Statistical Software Manual, I & II.* Berkeley, Calif: University of California Press; 1988.
16. Kleinbaum DG, Kupper LL, Muller KE. *Applied Regression Analysis and Other Multivariable Methods.* Boston, Mass: PWS-Kent Publishing Co; 1988:260–296.
17. Martone WJ, Kaufman AF. Leptospirosis in humans in the United States, 1974–1978. *J Infect Dis.* 1979;140:1020–1022.
18. Detmers RY, Frank RR, Thierman AB, Demers PA. Exposure to *Leptospira icterohemorrhagia* in inner-city and suburban children: a serologic comparison. *J Fam Pract.* 1983;17:1007–1011.
19. McClain JB. Leptospirosis. In: Wyngaarden JB, Smith LH, eds. *Cecil Textbook of Medicine.* Philadelphia, Pa: WB Saunders Co; 1988:1730–1732.
20. Cacciapuoti B, Vellucci A, Ciceroni L, Pinto A, Taggi F. Prevalence of leptospirosis in man: pilot study. *Eur J Epidemiol.* 1987;3:137–142.
21. Gale DA, Everard COR, Carrington DG, Everard JD. Leptospiral antibodies in patients from a Barbadian general practice. *Eur J Epidemiol.* 1990;6:150–155.
22. Cacciapuoti B, Ciceroni L, Maffei C, et al. A waterborne outbreak of leptospirosis. *Am J Epidemiol.* 1987;126:535–545.
23. Schmidt DR, Winn RE, Keefe TJ. Leptospirosis; epidemiological features of a sporadic case. *Arch Intern Med.* 1989;149:1878–1880.
24. Demers RY, Frank R, Demers P, Clay M. Leptospiral exposure in Detroit rodent control workers. *Am J Public Health.* 1985;76:1090–1091.
25. Childs JE. Urban cats: their demography, population density and owner characteristics in Baltimore, Maryland. *Anthrozoös.* 1990;3:234–244.
26. Teutsch SM, Juvanek DD, Sulzer A, Dubey JP, Sikes RK. Epidemic toxoplasmosis associated with infected cats. *N Engl J Med.* 1979;300:695–699.
27. Kaufmann AF, Mann JM, Gardiner TM, et al. Public health implications of plague in domestic cats. *J Am Vet Med Assoc.* 1981;179:875–878.

Characteristics Related to Elderly Persons' Not Eating for 1 or More Days: Implications for Meal Programs

Edward A. Frongillo, Jr, PhD, Barbara S. Rauschenbach, PhD, Daphne A. Roe, MD, and David F. Williamson, PhD

ABSTRACT

We examined how social, economic, location, health, and food need characteristics are related to elderly persons' not eating for 1 or more days. The following variables were positively related to not eating: ethnicity, location, receipt of Medicaid, living alone, health problems, mobility, age less than 80 years, cancer, nausea, difficulty swallowing, diarrhea, loss of appetite, and receipt of food from a food pantry. These results have implications for allocating meal program funds, screening clients, and monitoring whether clients eat regularly. (*Am J Public Health.* 1992;82:600–602)

Introduction

The demand for federally funded nutritional meals programs for the elderly has increased since 1972, when such programs were first authorized.[1] Persons 60 years or older and their spouses are eligible; priority is given to persons with incomes below the poverty line, to minority persons, and to the physically or mentally disabled or socially isolated. To reach those in greatest need, most agencies screen applicants for age, mobility, current illness and disability, and social isolation. However, few studies are available on characteristics of the elderly that are related to poor nutritional status, and these studies either did not explain variation in status or examined a limited set of explanatory variables.[2–15] In this paper we examine how economic, location, social, health, and food needs characteristics affect elderly persons' not eating for 1 or more days (Figure 1).

Methods

Meal program staff in New York City (NYC) and 31 New York State counties interviewed 4017 new home-delivered meal (HDM) clients in their homes from 1984 to 1986 before they began receiving meals. In 1985 and 1986, a self-administered survey assessed 2648 congregate meal (CM) clients already receiving meals.

The variable not eating for 1 or more days was modeled using logistic regression.[16] Dummy coding was used for most variables, the reference being those without the characteristic.[17] Receipt of Medicaid indicated economic status because poverty status had many missing values. Needing help with food shopping was omitted because it was closely related to needing help with food preparation. The reference for ethnicity, living in NYC, and their interaction was White living outside of NYC. The reference for sex, living alone, and their interaction was female not living alone. Number of health problems quantified diseases other than cancer. HDM and CM clients were analyzed separately because the two groups differed in health status, mobility, and whether they were receiving meals. Model fit was measured by area under the receiver operating characteristic (ROC) curve.[18]

Additional variables were available for 1985 and 1986 HDM clients: education, receipt of Supplemental Security Income, receipt of food from a food pantry, use of commodity foods, loss of appetite, and having diarrhea, nausea, difficulty swallowing, depression, constipation, incontinence, indigestion, dental problems, or impaired vision. The best extended model, of several fitted using these additional variables, was compared with the original model after it was refitted to the 1985 and 1986 subsample.

Results

The percentage of clients not eating for 1 or more days was 3.4% for CM and 17.5% for HDM clients. The variables that were important in explaining not eating were similar for the HDM and CM clients (Tables 1 and 2). The positively related variables were as follows: White living in NYC, minority living outside NYC, minority living in NYC, receiving Medicaid, male living alone, number of health problems, and being mobile only with aid. Needing help with food preparation was also found to be important, but its addition to the models made mobility no longer important and so it was excluded. Age 80

Edward A. Frongillo, Barbara S. Rauschenbach, and Daphne A. Roe are with the Division of Nutritional Sciences, Cornell University, Ithaca, NY. David F. Williamson is with the Nutrition Division, Center for Chronic Disease and Health Prevention, Centers for Disease Control, Atlanta, Ga. At the time of this study, he was with the Bureau of Nutrition, New York State Department of Health, Albany, NY.

Requests for reprints should be sent to Edward A. Frongillo, Jr, PhD, Division of Nutritional Sciences, B36 Savage Hall, Cornell University, Ithaca, NY 14853.

This paper was submitted to the Journal August 2, 1990, and accepted with revisions July 18, 1991.

years or older was negatively related. Having cancer and being chair- or bed-bound were positively related only for HDM clients; only 2% of CM clients had cancer and none of them were chair- or bed-bound. The effect of living alone differed with sex and with client group; the effect of ethnicity differed with location and client group. The magnitudes of the regression coefficients for the seven variables common to both models were smaller in the model for HDM clients. The ROC curve area was 0.82 for CM clients and 0.66 for HDM clients.

The best extended model for the 1985 and 1986 HDM clients included five additional variables: nausea, difficulty swallowing, diarrhea, loss of appetite, and receiving food from a food pantry (Table 3). The ROC curve area for the extended model (0.73) was larger than that for the original model refitted to the 1985 and 1986 subsample (0.66), indicating a better fit for the extended model.

TABLE 1—Variables Explaining Not Eating for 1 or More Days among Home-Delivered Meal Clients (n = 3353)

	Percentage of Sample	Regression Coefficient	Odds Ratio	95% CI for Odds Ratio	
Constant		−2.04			
Minority living outside NYC	3	0.27	1.31	0.77	2.24
White living in NYC	16	0.37	1.45	1.13	1.86
Minority living in NYC	6	0.84	2.33	1.67	3.24
Receiving Medicaid	21	0.45	1.58	1.27	1.96
Age 80 years or older	41	−0.32	0.73	0.59	0.89
Female living alone	44	0.03	1.03	0.82	1.30
Male not living alone	17	−0.54	0.58	0.42	0.81
Male living alone	15	0.23	1.26	0.94	1.69
Number of health problems[a]	1.3	0.09	1.21	1.02	1.43
Having cancer	8	0.73	2.08	1.55	2.79
Being mobile with aids	46	0.16	1.17	0.96	1.43
Being chair- or bed-bound	6	0.74	2.10	1.46	3.01

Note. CI = confidence interval; NYC = New York City. The likelihood ratio chi-square statistic for comparison with the model having only the constant was 143.68 (12 *df*). The area under the ROC curve was 0.66. [a]The mean is shown (range, 0 to 4). The odds ratio and 95% CI were calculated for two times the regression coefficient; the factor of two corresponded to two standard deviations. This calculation allowed for a more meaningful comparison of the odds ratio with that of the other variables.

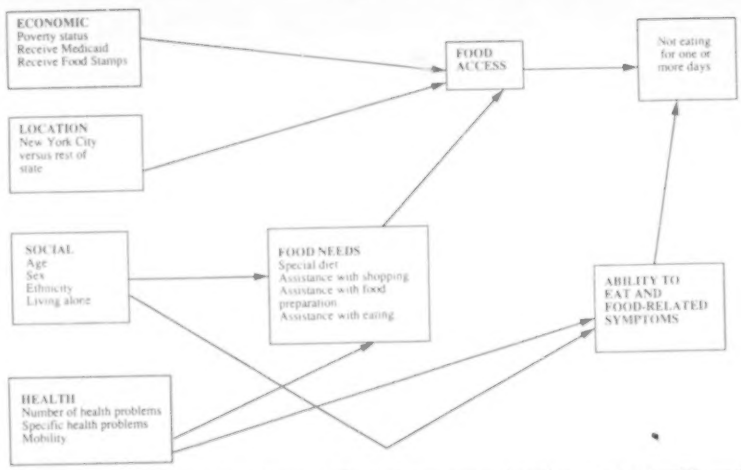


FIGURE 1—Conceptual map for the effect of economic, location, social, health, and food needs characteristics, through their effect on both access to food and the ability and desire to eat, on not eating for 1 or more days.

Discussion

HDM clients had more characteristics related to not eating than did CM clients, but this explains only some of the group difference in the percentage of those who did not eat; the larger constant in the model for the HDM clients indicates that their odds of not eating were greater, even when they had none of the characteristics included in the model. The model for HDM clients did not explain not eating as well as did the model for CM clients. Some characteristic(s) other than those included affected whether the HDM clients ate regularly. Indeed, the extended model for HDM clients was substantially better at explaining who did not eat for 1 or more days.

By increasing clients' access to food, participation in the CM program may have reduced the odds of their not eating, because these clients were already receiving meals. The home-delivered meal program could be expected to have a similar effect on clients once they began receiving meals.

For HDM clients, the finding that living alone was positively related to not eating among men, but not among women, is similar to that previously reported with national survey data.[13-15] However, just as this interaction did not occur for both groups of clients in our study, it occurred in only some of the previous studies and for only some measures of eating behavior.

The lower odds of not eating among persons 80 years or older may seem counterintuitive. Persons who live to age 80 may have certain characteristics, not included in the models, that allow them to eat more regularly.

The characteristics we found to be related to not eating can help determine allocation of meal program funds. Counties with more elderly residents having characteristics related to not eating for 1 or more days should be considered to be in greater need. These characteristics can also be useful for developing screening criteria to ensure that those most likely to not eat regularly are enrolled in the programs, and for nutritional monitoring of clients after enrollment. Elderly persons who have many of the characteristics at the time of enrollment may be monitored to ensure that they do not subsequently go without food.

Some CM clients did not eat for 1 or more days despite their participation in the CM program. Most programs provide meals 5 days per week, but these clients may need congregate meals all 7 days. Among HDM clients, those with nutrition-related health problems may need nutritional counseling, medical attention, and encouragement—in addition to meals—to ensure adequate nutrition. □

TABLE 2—Variables Explaining Not Eating for 1 or More Days among Congregate Meal Clients (n = 2240)

	Percentage of Sample	Regression Coefficient	Odds Ratio	95% CI for Odds Ratio	
Constant		−5.50			
Minority living outside NYC	6	1.11	3.05	1.47	6.32
White living in NYC	14	1.41	4.08	2.28	7.29
Minority living in NYC	14	0.85	2.34	1.25	4.40
Receiving Medicaid	20	1.16	3.20	1.98	5.18
Age 80 years or older	18	−0.99	0.37	0.17	0.80
Female living alone	42	0.94	2.56	1.17	5.59
Male not living alone	19	0.62	1.86	0.72	4.86
Male living alone	16	1.58	4.85	2.13	11.07
Number of health problems[a]	1.3	0.29	1.30	1.19	2.72
Being mobile with aids	18	0.49	1.63	0.98	2.72

Note. CI = confidence interval; NYC = New York City. The likelihood ratio chi-square statistic for comparison with the model having only the constant was 125.43 (10 *df*). The area under the ROC curve was 0.82. [a]The mean is shown (range, 0 to 4). The odds ratio and 95% CI were calculated for two times the regression coefficient; the factor of two corresponded to two standard deviations. This calculation allowed for a more meaningful comparison of the odds ratio with that of the other variables.

TABLE 3—Variables Explaining Not Eating for 1 or More Days among Home-Delivered Meal Clients in Extended Model for 1985 to 1986 survey (n = 1431)

	Percentage of Sample	Regression Coefficient	Odds Ratio	95% CI for Odds Ratio	
Constant		−2.15			
Minority living outside NYC	2	1.33	3.77	1.70	8.37
White living in NYC	9	0.56	1.75	1.07	2.84
Minority living in NYC	3	1.16	3.19	1.59	6.43
Receiving Medicaid	21	0.47	1.60	1.15	2.25
Age 80 years or older	42	−0.49	0.61	0.45	0.84
Female living alone	45	−0.10	0.90	0.63	1.30
Male not living alone	16	−0.28	0.75	0.46	1.23
Male living alone	16	0.34	1.41	0.89	2.22
Number of health problems[a]	1.9	−0.04	0.92	0.71	1.20
Having cancer	11	0.41	1.51	1.00	2.28
Being mobile with aids	41	−0.03	0.97	0.71	1.32
Being chair- or bed-bound	5	0.52	1.68	0.94	3.01
Having nausea	16	0.36	1.44	0.99	2.09
Having difficulty swallowing	12	0.43	1.54	1.03	2.30
Having diarrhea	12	0.41	1.51	1.01	2.26
Loss of appetite	34	0.91	2.48	1.83	3.37
Receiving food from food pantry	5	0.69	2.00	1.14	3.49

Note. CI = confidence interval; NYC = New York City. The likelihood ratio chi-square statistic for comparison with the model having only the constant was 153.36 (17 *df*). The likelihood ratio chi-square statistic comparing this extended model with the original model for the 1985–1986 clients was 78.09 (5 *df*). The area under the ROC curve was 0.73. [a]The mean is shown (range, 0 to 4). The odds ratio and 95% CI were calculated for two times the regression coefficient; the factor of two corresponded to two standard deviations. This calculation allowed for a more meaningful comparison of the odds ratio with that of the other variables.

Acknowledgments

This research was funded in part by the New York State Department of Health. We are grateful for the interest and help of Janice Dodds, former Director of Nutrition Surveillance, New York State Department of Health, and Dorothy Diggins, New York State Office for Aging, who were instrumental in the initiation and support of this study. We especially thank the dedicated people in the county and local program offices for their fine work in collecting the data.

References

1. Kirschner Associates. An evaluation of the nutrition services for the elderly, vol. 3: descriptive report. Washington, DC: Administration on Aging, US Dept of Health and Human Services; 1983. US Dept of Health and Human Services publication OHDS 83-20917.
2. Exton-Smith AN. Nutritional status: diagnosis and prevention of malnutrition. In: Exton-Smith AN, ed. *Metabolic and Nutritional Disorders in the Elderly.* Bristol, England: John Wright; 1980.
3. Davies L. Nutrition and the elderly: identifying those at risk. *Proc Nutr Soc.* 1984;43:295–302.
4. Goodwin JS, Goodwin JM, Garry PJ. Association between nutritional status and cognitive functioning in a healthy elderly population. *JAMA.* 1983;249:2917–2921.
5. Morris JN, Sherwood S, Mor V. An assessment tool for use in identifying functionally vulnerable persons in the community. *Gerontologist.* 1984;24:373–379.
6. Caird FI, Judge TG, Macleod C. Pointers to possible malnutrition in the elderly at home. *Gerontol Clin.* 1975;17:47–54.
7. McIntosh WA, Shifflett PA. Influence of social support systems on dietary intake of the elderly. *J Nutr Elderly.* 1984;4(1): 5–17.
8. Rawson IG, Weinberg EI, Heroid JA, Holtz J. Nutrition of rural elderly in southwestern Pennsylvania. *Gerontologist.* 1978;18:24–29.
9. Vaughan LA, Manore MM. Dietary patterns and nutritional status of low income, free-living elderly. *Food Nutr News.* 1988;60:27–36.
10. Steele MF, Bryan JD. Dietary intake of homebound elderly recipients and nonrecipients of home-delivered meals. *J Nutr Elderly.* 1986;5(2):23–34.
11. Caliendo MA, Batcher M. Factors influencing the dietary status of participants in the National Nutrition Program for the Elderly. *J Nutr Elderly.* 1980;1(1):41–53.
12. Betts NM, Vivian VM. Factors related to the dietary adequacy of noninstitutionalized elderly. *J Nutr Elderly.* 1985;4(4):3–14.
13. Davis MA, Randall E, Forthofer RN, Lee ES, Margen S. Living arrangements and dietary patterns of older adults in the United States. *J Gerontol.* 1985;40:434–442.
14. Davis MA, Murphy SP, Neuhaus JM. Living arrangements and eating behaviors of older adults in the United States. *J Gerontol.* 1988;43:S96–S98.
15. Ryan AS, Martinez GA, Wysong JL, Davis MA. Dietary patterns of older adults in the United States, NHANES II 1976–1980. *Am J Hum Biol.* 1989;1:321–330.
16. SAS Institute. SAS user's guide: basics and statistics. Cary, NC: SAS Institute; 1982.
17. Pedhazur EJ. Multiple regression in behavioral research. New York, NY: Holt, Rinehart & Winston; 1982.
18. Swets JA. Measuring the accuracy of diagnostic systems. *Science.* 1988;240:1285–1293.

The South Karelia Air Pollution Study: Acute Health Effects of Malodorous Sulfur Air Pollutants Released by a Pulp Mill

Tari Haahtela, MD, Olli Marttila, MSc, Vesa Vilkka, MD, Paavo Jäppinen, MD, and Jouni J. K. Jaakkola, MD

ABSTRACT

We evaluated the acute health effects of a strong emission of malodorous sulfur compounds released from a pulp mill in South Karelia, Finland. The 24-hour ambient air concentrations of hydrogen sulfide for the two emission days were 35 and 43 $\mu g/m^3$ (maximum 4-hour 135 $\mu g/m^3$). A questionnaire was distributed after the high exposure and later after a low exposure period to 29 households with 75 subjects living in the nearby community. During the high exposure, 63% of the respondents reported experience of at least one symptom compared to 26% during the reference period. Every third participant reported difficulties in breathing. In the 45 subjects responding to both questionnaires more eye, respiratory and neuropsychological symptoms occurred during the exposure compared to the reference period. The strong malodorous emission from a pulp mill caused an alarming amount of adverse effects in the exposed population. (*Am J Public Health.* 1992;82:603–605)

Introduction

Large populations in northern Europe and North America are often or constantly exposed to malodorous sulfur air pollutants such as hydrogen sulfide (H_2S), methyl mercaptan (CH_3SH), and methyl sulfides produced by the paper and pulp industry. The knowledge of the adverse effects of these pollutants is sparse, although people living in the neighborhood affected commonly complain of discomfort and express concern about the health effects.

A cluster of paper mills producing wood pulp is located in South Karelia, in southeastern Finland. The two sulfate pulp mills, located in the small city of Imatra (33 000 inhabitants), were, in 1989, estimated to emit annually 2800 tons of H_2S into the air. The annual mean (2 $\mu g/m^3$) concentration of sulfur dioxide (SO_2) is low because of the use of natural gas as the source of industrial energy.

During 2 days in September 1987, a strong emission, with H_2S as the major component and mesityloxide as a minor component, was released from one of the two pulp mills, exposing the local population to intensive catty odor for the 2 days. We evaluated the acute health effects of the emission by comparing the occurrence of symptoms during the high-exposure period and 4 months later during a low-exposure period.

Material and Methods

During the introduction of a new processing technique, a strong malodorous emission was released from the pulp mill using the sulfate method for production, and increased concentrations of H_2S were measured for 2 days at our monitoring station located in a nearby community about 1 km southeast of the source of exposure. The H_2S exposure was measured by absorbing the air sample to cadmium hydroxide.[1] The 4-hour concentrations of H_2S during the high exposure were, on the average, 4 to 5 times higher, and a maximum of 20 times higher, than the levels before and after the peak emission, during which there was inversion of the air and no wind (Figure 1). The highest 4-hour concentration of H_2S, 135 $\mu g/m^3$, was measured at 2 AM. The 24-hour averages for the 2 days were 35 and 43 $\mu g/m^3$.

During the peak emission of malodorous sulfur compounds, the mean 1-hour concentration of SO_2 was only 3 $\mu g/m^3$. It was not possible to measure mesityloxide in the ambient air. Mesityloxide is produced inside pulp mills under the same kinds of conditions that occurred during the peak emission. The presence of mesityloxide was recognized in the air by its characteristic catty odor.

A reference period of 2 days was chosen about 4 months later, in January 1988. During the 2 reference days no odor was sensed, and the concentration of H_2S was low (0.1 to 3.5 $\mu g/m^3$ for 4 hours). The SO_2 concentration was low, 3 $\mu g/m^3$, as during the high-exposure period.

An open-ended questionnaire was distributed 10 days after the high exposure to all 29 households of the nearby community; 66 (88%) of a possible 75 subjects responded. Six subjects were excluded because of missing information, leaving 60 subjects: 44 adults (15 men and 29 women; mean age = 45 years and range = 17 to

Tari Haahtela is with the Department of Allergic Diseases, Helsinki University Central Hospital, Helsinki, Finland. Olli Marttila is with the South Karelia Allergy and Environment Institute, Joutseno, Finland. Vesa Vilkka is with the Department of Pulmonary Diseases, Tiuru Hospital, Joutseno, Finland. Paavo Jäppinen is with the Vuoksenlaakson Occupational Health Centre, Imatra, Finland. Jouni J. K. Jaakola is with the Department of Public Health, University of Helsinki, Helsinki, Finland.

Requests for reprints should be sent to Tari Haahtela, MD, Department of Allergic Diseases, Helsinki University Central Hospital, Meilahdentie 2, 00250 Helsinki, Finland.

This paper was submitted to the Journal November 26, 1990, and accepted with revisions July 9, 1991.

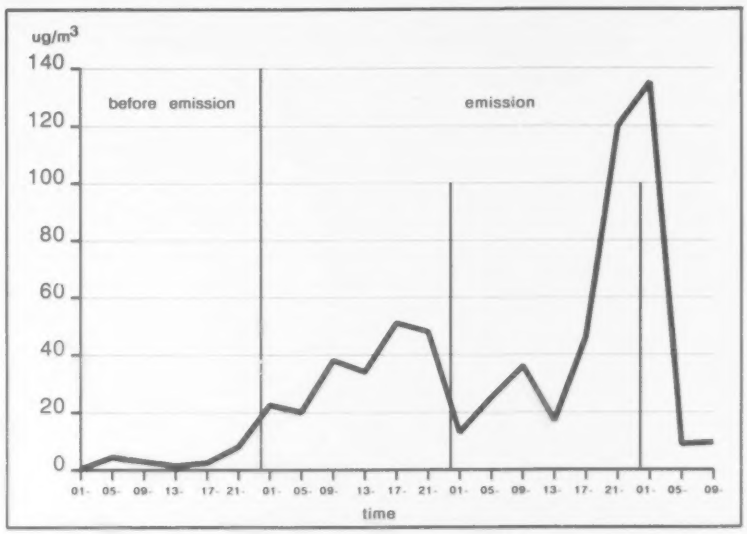


FIGURE 1—Four-hour concentrations of hydrogen sulfide before and during the emission from the first day at 1 AM to the fourth day at 9 AM at the monitoring station located about 1 km from the pulp mill.

TABLE 1—Prevalence of Symptoms in Subjects Who Responded after the High-Exposure Period (n = 60) and after the Reference Period (n = 66)

	High-Exposure Period		Reference Period	
	n	%	n	%
Eye symptoms	13	22	1	2***
Males	5	21	0	0
Females	8	22	1	3
Nasal symptoms	5	8	13	20
Males	2	8	5	16
Females	3	8	8	24
Cough or pharyngeal irritation	9	15	3	5
Males	5	21	1	3
Females	4	11	2	6
Breathlessness	21	35	1	2***
Males	7	29	0	0
Females	14	39	1	3
Nausea	14	23	3	5**
Males	8	33	1	3
Females	6	17	2	6
Headache	14	23	5	8*
Males	5	21	1	3
Females	9	25	4	12
Mental symptoms[a]	6	10	0	0*
Males	1	4	0	0
Females	5	14	0	0

Note. Differences between prevalences were obtained using a χ^2-test.
[a]Depression, anxiety.
*$p < .05$.
**$p < .01$.
***$p < .001$.

85) and 16 children (9 boys and 7 girls; mean age = 7 years and range = 1 to 15). The respondents were asked to describe the symptoms they had experienced during the days of peak exposure, without mention of the exceptionally high emission of pollutants or what kinds of symptoms the emission may generate. Parents responded on behalf of their children.

An identical questionnaire was distributed to the same households 4 months later. The link of this reference questionnaire to the first inquiry was not revealed. The response rate was 79%, no data had to be excluded, and 66 persons were included in the analyses. The effect of confounding was controlled by studying the intraindividual difference in the symptoms in the subpopulation of 45 subjects who responded to both questionnaires. The statistical significance of the difference in the occurrence of the symptoms during the high-exposure and reference periods was assessed by a χ^2-test and McNemar's test.[2]

Results

The unadjusted prevalences of the different symptoms in the two population samples are given in Table 1. Large proportions of the participants reported breathlessness (35%) and mental symptoms (10%) during the high-exposure period, while no such symptoms were indicated during the reference period. The prevalence of eye symptoms (22% vs 2%), cough or pharyngeal irritation (15% vs 5%), and nausea (14% vs 3%) were also significantly larger during the high exposure. However, nasal symptoms (8% vs 20%) were more common during the reference period than the high-exposure period, although the difference was not significant. Altogether, 63% of the subjects reported experience of at least one symptom during the high-exposure period compared with 26% during the reference period.

The presence of symptoms during the exposure and reference periods in the 45 subjects who responded to both questionnaires is shown in Table 2. Eye symptoms, cough or pharyngeal irritation, breathlessness, nausea, and headache were experienced more often during the exposure period than during the reference period. The difference was significant for breathlessness.

Discussion

The occurrence of the symptoms during the high-exposure period was large: Every third participant reported difficulties in breathing, and every fourth participant indicated irritation of eyes, headache, and nausea. Mental symptoms were also common. The prevalence of these symptoms was significantly larger than those during the reference period in the comparison of the two samples. The direction of the differences was similar in the intraindividual comparison of those who responded to both questionnaires, al-

though significance was not reached in all of the symptoms because of the smaller number of subjects than in the two-sample comparison. Interestingly, more subjects reported nasal symptoms during the reference period than during the high-exposure period. A possible explanation is that breathing difficulties due to exposure were so prominent that attention was not paid to mild nasal symptoms. The highest 4-hour concentration of H_2S was 135 $\mu g/m^3$, which is nine times higher than the guideline in Canada (15 $\mu g/m^3$). There are no guidelines for ambient air H_2S in Finland.

The observed symptoms correspond to physiological effects of acute exposure of H_2S suggesting direct irritative effect on mucous membranes and eye conjunctivae, although in far smaller concentrations than described earlier.[3] Previously, we have observed an increased amount of eye and nasal symptoms and cough among subjects living in the most polluted area of Imatra.[4]

According to the World Health Organization (WHO) guideline values for H_2S, the highest acceptable half-hour concentration is 7 $\mu g/m^3$ for odor nuisance, along with a daily mean of 150 $\mu g/m^3$ for health hazards.[5] During the period of high exposure, the highest measured 4-hour concentration (135 $\mu g/m^3$) and 24-hour averages (35 and 42 $\mu g/m^3$) were well below the recommendation. These kind of emissions are rather commonplace and are always possible wherever wood pulp production takes place. It seems likely that the WHO guideline value of 150 $\mu g/m^3$ for H_2S is too high and does not provide prevention from adverse health effects. Furthermore, the presence of other pollutants should be taken into account, because H_2S is rarely the only hazardous component of air pollution. □

TABLE 2—Occurrence of Symptoms during the High-Exposure and Reference Periods in the 45 Subjects Who Responded to Both Questionnaires

	Exposure Period Only, No.	Reference Period Only, No.	Both Periods, No.	No Symptoms, No.	Significance[a]	
					t	P
Eye symptoms	5	1	0	39	1.63	NS
Nasal symptoms	3	11	1	30	2.14	<.05
Cough or pharyngeal irritation	5	1	0	39	1.63	NS
Breathlessness	16	0	1	28	4.00	<.001
Nausea	6	2	1	36	1.41	NS
Headache	8	2	2	33	1.90	NS
Mental symptoms[b]	2	0	0	43	1.41	NS

Note. NS = not significant.
[a]McNemar's test of significance.
[b]Depression, anxiety.

Acknowledgments

This study was supported by the Academy of Finland and the Ministry of the Environment.

References

1. Air Quality Determination of the mass concentration of hydrogen sulfide in ambient air. Methylene blue spectrophotometric method. International Organization for Standardization ISO/DIS 6769, 1981.
2. Armitage P. *Statistical Methods in Medical Research.* Oxford and Edinburgh, England: Blackwell Scientific Publications; 1971.
3. Beauchamp RO, Bus JS, Popp JA, Boreiko CJ, Andjelkovich DA. A critical review of the literature on hydrogen sulfide toxicity. *CRC Crit Rev Toxicol.* 1984;13:25–97.
4. Jaakkola JJK, Vilkka V, Marttila O, Jäppinen P, Haahtela T. The South Karelia air pollution study. The effects of malodorous sulfur compounds from pulp mills on respiratory and other symptoms. *Am Rev Respir Dis.* 1990;142:1344–1350.
5. World Health Organization. *Air Quality Guidelines for Europe.* Denmark: World Health Organization; 1987. WHO regional publications. European series No 23. Copenhagen.

The Epidemiology of Low Back Pain in an Adolescent Population

ABSTRACT

We assessed the prevalence of low back pain (LBP) in a cohort of 1242 adolescents (aged 11 through 17) currently participating in a 4-year prospective study of medically treated injuries. Overall, 30.4% of the adolescents reported LBP. The impact of LBP in adolescents was considerable, with one third resulting in restricted activity and 7.3% seeking medical attention. Life-table analysis demonstrated that by age 15, the prevalence of LBP increased to 36%. There were few differences by gender or race. These results suggest that LBP in adolescents is a serious public health problem. (*Am J Public Health.* 1992;82:606–608)

Todd L. Olsen, MS, MPH, Robyn L. Anderson, MS, MPH, Stephen R. Dearwater, MS, Andrea M. Kriska, PhD, Jane A. Cauley, DrPH, Deborah J. Aaron, MS, and Ronald E. LaPorte, PhD

Introduction

Recent figures indicate that 70% to 80% of the general population will suffer from low back pain (LBP) at some point in their lives.[1–3] Research on LBP has focused primarily on occupational studies in adult populations. Little is known about the development of LBP in children. The epidemiology of LBP in adolescence is of particular importance, as LBP may occur initially during this time and may foreshadow the subsequent severe, chronic morbidity seen in adulthood.

There are only limited data on the prevalence of LBP in adolescents, probably because the disorder in this age group has been thought to be rare. Turner and colleagues[4] found that less than 2% of all children under age 15 presented with LBP at an orthopedic practice. Additionally, Grantham,[5] a school doctor in Northamptonshire, United Kingdom, found that 11.5% of his adolescent male students were seen because of LBP. In contrast to these clinical studies, two school-based studies from Europe have shown that backache may be quite common in this age group. Specifically, Balague et al.[6] found that 27% of 1715 Swiss school pupils complained of LBP as assessed from a validated questionnaire. A similar prevalence of LBP (26%) was found in English school pupils studied by Fairbank et al.[7]

From these few studies, the prevalence of LBP in adolescents in the United States is not clear, nor are the demographic factors associated with the disorder. In addition, there is no information concerning when the first symptoms of LBP appear in populations. Therefore, the proposed research evaluated for the first time in the United States the epidemiology of LBP in a biracial population of adolescents.

Population and Methods

Data on LBP were collected as part of an ongoing prospective population-based investigation of the epidemiology of adolescent injuries. The population undergoing injury surveillance consisted of 1246 (89%) of the 1400 seventh through ninth graders from an urban school district in Allegheny County, Pa. The demographic characteristics of the school district were virtually identical to those of the county.

LBP was ascertained using a modified questionnaire of Fairbank et al., a copy of which is included in the Appendix. The questionnaire was designed to identify any occurrence of LBP in adolescents, its severity, age of onset, recurrence, activity limitation, and medical treatment. The reliability of the survey has been found to be high, with an 84% consistency rate over 6 months.[2]

Results

Ninety-nine percent (1242/1246) of the students participating in the study completed the LBP questionnaire. Table 1 presents the demographic characteristics of the sample. There were 641 male and 601 female subjects 11 through 17 years of age, with an average age of 13.6 years. This represented students in grades seven (32.7%), eight (33.0%), and nine (34.3%).

A surprisingly high percentage, 30.4% (95% confidence interval [CI] = 28.7, 32.2) of the children reported ever experiencing low back pain (Question 1). Twenty-two percent (95% CI = 20.5, 23.0) of the children reported experiencing LBP within the previous 12 months (Question 3). Seven percent (95% CI = 6.9, 7.7) of those who reported LBP sought medical attention at least once (Question 5). There was a greater prevalence of LBP leading to medical attention for girls, 8.0% (95% CI = 7.4, 8.7) vs boys,

Todd L. Olsen, Robyn L. Anderson, Andrea Kriska, Jane A. Cauley, Deborah Aaron, and Ronald E. LaPorte are with the Department of Epidemiology, University of Pittsburgh, Pittsburgh, Pa. Stephen Dearwater is with the Sports Medicine Institute, University of Pittsburgh Medical Center, Pittsburgh, Pa.

Requests for reprints should be sent to Todd L. Olsen, Department of Epidemiology, A529 Crabtree Hall, University of Pittsburgh, Pittsburgh, PA 15261.

This paper was submitted to the Journal January 2, 1991, and accepted with revisions June 19, 1991.

6.7% (95% CI = 6.2, 7.2). In addition, 8.8% (95% CI = 8.3, 9.3), of the students indicated that the back pain was sufficiently severe to force them to miss school or keep them from playing sports (Question 4).

Figure 1 presents the prevalence of LBP by age and race. The only difference in LBP prevalence when examined by race occurred at age 15 where Blacks had significantly higher rates than Whites (47% vs 31%, $P < .05$).

Life-table analysis, as seen in Figure 2, demonstrated that by age 15 the likelihood of this cohort reporting LBP was 36%. There were no differences between boys and girls at any age. It should also be noted that LBP was rare under age 10. This finding can, however, be potentially related to subject recall. Low back pain was additionally analyzed by the hazard function in Figure 3, which can be interpreted as an age-at-onset incidence curve. There was a markedly accelerating incidence of LBP beginning at about age 10, with a continuously increasing risk through adolescence. By age 15, the risk of reporting the first occurrence of LBP was twice that of 12-year-olds, suggesting that the yearly risk of LBP increases dramatically during the teenage years.

Discussion

These results suggest that 1 in every 3 adolescents report experiencing LBP in their lifetime. If these results are extrapolated to the oldest age group, by age 15, 36% of the children would be expected to report the occurrence of LBP. Thus, LBP in adolescent children is a serious public health problem.

Our results are consistent with those seen in European population-based studies. Specifically, Fairbank et al.[7] and Balague et al.[6] found LBP rates of 26% and 27%, respectively.

Approximately one third of the self reported LBP was associated with activity limitations, with one in four of those suffering from LBP seeking medical treatment. Although this represents 3% and 2% of our adolescent cohort, respectively, the economic and public health implications are substantial when one considers the young age and recurrence potential. In a compreshensive review of the literature, Kelsey et al.[3] report that the first episode of LBP occurs most often between the ages of 20 and 40 years; however, these data have been based on recall of adults sometimes decades after the initial episode. Furthermore, costs associated with adult LBP presently exceed estimates of $16 billion.[8] Our results suggest that the initial onset and subsequent cost of LBP, may, in fact, occur early in youth.

There was a similar proportion of reported LBP in boys (30.7%) and girls (30.0%). As presented, it was unlikely that the LBP in girls was exclusively related to menstrual cycle because of the lack of sex differences. It is possible that potential determinants for LBP are thus similar for boys and girls.

The only observed racial differences in LBP were observed in the oldest adolescents. By age 15, blacks reported significantly higher prevalence rates of LBP than whites. There has been no other population based studies in children where it has been possible to examine race-specific prevalence rates.

Our findings, in accordance with other research, suggest that LBP in adolescents is potentially a very serious prob-

TABLE 1—Population Characteristics

Demographics	Cohort (n = 1242) n	%
Sex		
Male	641	51.6
Female	601	48.4
Race		
White	900	72.5
Black	303	24.4
Other	32	2.5
Age		
11	1	0.1
12	202	16.3
13	369	29.7
14	395	31.8
15	233	18.8
16	38	3.1
17	4	0.3
Grade		
7	406	32.7
8	410	33.0
9	426	34.3

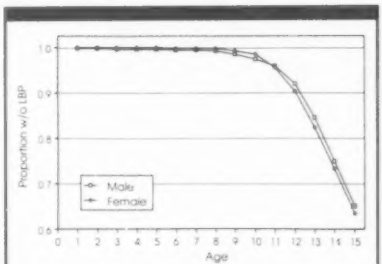


FIGURE 2—Life-table analysis of low back pain (LBP) for boys and girls.

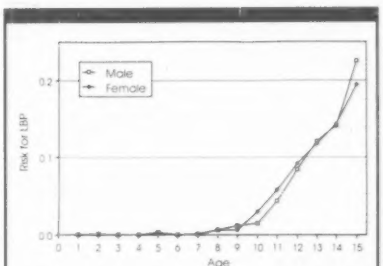


FIGURE 3—Instantaneous hazard function for boys and girls.

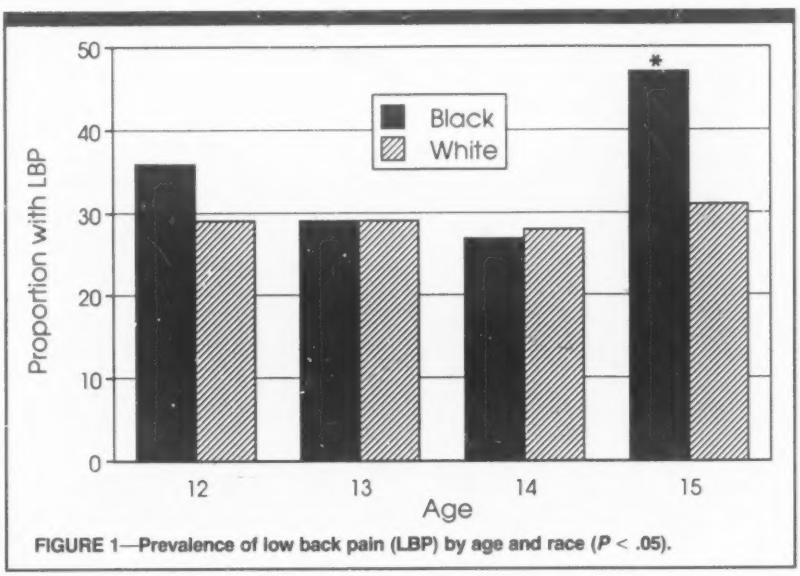


FIGURE 1—Prevalence of low back pain (LBP) by age and race ($P < .05$).

APPENDIX—Low Back Pain Assessment by Descriptive Questionnaire

1. Have you ever had pain or other trouble with the lower part of your back? (yes/no)
2. How old were you the first time that you had trouble with your back? (__ years old)
3. Have you had pain or other trouble with the lower part of your back within the last 12 months? (yes/no)
4. Has your back pain ever stopped you from going to school or playing sports? (yes/no)
5. Have you ever seen your doctor about back pain? (yes/no)

lem. Future research in adolescents must address the etiology, associated risk factors, and severity and the degree to which these predict subsequent recurrence of LBP. This may help to identify those adolescents who are at greatest risk for developing debilitating LBP in adulthood so that preventive measures can be undertaken to reduce its occurrence. □

Acknowledgments

This study was supported by the National Institute of Arthritis, Musculoskeletal and Skin Disease Grant 5 R01 AR39541-02.

References

1. Frymoyer JW, Pope MH, Clement JH, et al. Risk factors in low-back pain. *J Bone Jt Surg.* 1983;65A:213–218.
2. Biering-Sorensen FA. A prospective study of low back pain in a general population. *Scand J Rehabil Med.* 1984;31:362–374.
3. Kelsey JL, Golden AL, Mundt DJ. Low back pain/prolapsed lumbar intervertebral disc. *Rheum Dis Clin North Am.* 1990; 16:699–715.
4. Turner PG, Green JH, Galasko CSB. Back pain in childhood. *Spine.* 1989;14:812–814.
5. Grantham VA. Backache in boys—a new problem. *Practitioner.* 1977;218:226–229.
6. Balague F, Dutoit G, Waldburger M. Low back pain in schoolchildren. *Scand J Rehabil Med.* 1988;20:175–179.
7. Fairbanks J, Pynsent PB, Poortvliet JA, et al. Influence of anthropometric factors and joint laxity in the incidence of adolescent back pain. *Spine.* 1984;9:461–464.
8. Frymoyer JW, Gorden SL. Research perspectives in low-back pain—report of a 1988 workshop. *Spine.* 1989;14:1364–1387.

Call for Abstracts on Chronic Disease Prevention and Control

The Centers for Disease Control, the Association of State and Territorial Health Officials, and the Association of State and Territorial Chronic Disease Program Directors announce the Seventh National Conference on Chronic Disease Prevention and Control to be held at the Salt Lake Hilton Hotel in Salt Lake City, Utah, October 21 to 23, 1992. The Utah Department of Health will host the conference, which is open to the public.

Abstracts are being solicited and will be considered for concurrent oral presentations or poster sessions dealing with (a) assessment and evaluation, (b) chronic disease epidemiology, and (c) program applications.

Deadline for submission is Friday, May 1, 1992. For more information, please contact Jack Friel, NCCD PHP, Mail Stop K-43, Centers for Disease Control, Atlanta, GA 30333. Tel: 404/488-5390; Fax: 404/488-5962.

American Journal of Public Health

May 1992, Vol. 82, No. 5 Established 1911

86389
028
SVC
ED KIANKA INFORMATION
CONGRESSIONAL HIGHWAY
4520 EAST-WEST
BETHESDA MD 20814

Editorials

663 Women's Health, Women's Lives, Women's Rights *H. Rodriguez-Trias*

664 Women and Children First: Towards a US Family Policy *R. Sidel*

Commentary

666 Folic Acid and Neural Tube Defect: Can't We Come to Closure? *W. C. Willett*

Featuring Women's Health

669 Barrier Contraceptives and Sexually Transmitted Diseases in Women: A Comparison of Female-Dependent Methods and Condoms *M. J. Rosenberg, A. J. Davidson, J.-H. Chen, F. N. Judson, and J. M. Douglas*

675 Pregnancy-Induced Hypertension in North Carolina, 1988 and 1989 *D. A. Savitz and J. Zhang*

680 The Transformation of American Midwifery: 1975 to 1988 *E. R. Declercq*

685 The Influence of Social and Political Violence on the Risk of Pregnancy Complications *B. C. Zapata, A. Rebolledo, E. Atalah, B. Newman, and M.-C. King*

691 A New Definition of Maternal Depletion Syndrome *A. Winkvist, K. M. Rasmussen, and J.-P. Habicht*

695 Sources of Fiber and Fat in Diets of US Women Aged 19 to 50: Implications for Nutrition Education and Policy *F. E. Thompson, M. F. Sowers, E. A. Frongillo, Jr, and B. J. Parpia*

719 Home Apnea Monitoring and Disruptions in Family Life: A Multidimensional Controlled Study *E. Ahmann, L. Wulff, and R. G. Meny*

723 Carrier Screening for Cystic Fibrosis: A Pilot Study of the Attitudes of Pregnant Women *J. R. Botkin and S. Alemagno*

726 A Cohort Study of the Impact of Perinatal Drug Use on Prematurity in an Inner-City Population *J. G. Feldman, H. L. Minkoff, S. McCalla, and M. Salwen*

729 The Impact of Mailing Psychoeducational Materials to Women with Abnormal Mammograms *C. Lerman, E. Ross, A. Boyce, P. McGovern Gorchov, R. McLaughlin, B. Rimer, and P. Engstrom*

731 Breast-feeding Patterns in a Rural Village in Giza, Egypt *I. A. Hakim and I. A. El-Ashmawy*

733 Access to Cancer Screening Services for Women *B. Kirkman-Liff and J. J. Kronenfeld*

736 Moderate Physical Activity and Bone Density among Perimenopausal Women *J. Zhang, P. J. Feldblum, and J. A. Fortney*

(continued page 651)

APHA
American Public Health Association

Featuring Women's Health

Contents

American Journal of Public Health

The Journal's one-page guidelines for authors, "What *AJPH* Authors Should Know," are printed in each issue. They can also be obtained from the Editorial Office.

May 1992, Vol. 82, No. 5 ISSN: 0090-0036 Established 1911

Articles

703 Overcoming the Absence of Socioeconomic Data in Medical Records: Validation and Application of a Census-Based Methodology *N. Krieger*

711 Ownership and Performance of Outpatient Substance Abuse Treatment Centers *J. R. C. Wheeler, H. Fadel, and T. A. D'Aunno*

Public Health Briefs

739 Medical Expenditures for Disability and Disabling Comorbidity *D. P. Rice and M. P. LaPlante*

742 Swimming-Associated Cryptosporidiosis *F. J. Sorvillo, K. Fujioka, B. Nahlen, M. P. Tormey, R. Kebabjian, and L. Mascola*

745 Newspapers: A Source for Injury Surveillance? *D. Y. Rainey and C. W. Runyan*

747 Response Comparability of Family and Staff Proxies for Nursing Home Residents *S. L. Tennstedt, K. M. Skinner, L. M. Sullivan, and J. B. McKinlay*

Other Journal Departments

652 Masthead

654 Advertisers' Index

656 What *AJPH* Authors Should Know

718 Errata

750 Association News: In Memoriam

755 Letters to the Editor

763 Book Corner

769 Job Opportunities

APHA
American Public Health Association

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.

APHA OFFICERS

Joyce C. Lashof, MD, *President*

C. William Keck, MD, MPH, *Immediate Past President*

Helen Rodriguez-Trias, MD, *President-Elect*

Arnold E. Greenberg, SM, *Vice-President (USA)*

Trevor Hancock, MB, MHSc, *Vice-President (Canada)*

Julio Teja Perez, MD, *Vice-President (Latin America)*

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Frances M. Veverka, RS, MPH, *Chair, Executive Board*

George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*

William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD

Frances M. Veverka, RS, MPH, (1992), *Chair*

Marilyn Aguirre-Molina, EdD, MS (1993)

Brian A. Dobrow, MPH (1992)

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Kristine M. Gebbie, BSN, MN, (1994), *Vice-Chair*

Alan R. Hinman, MD, MPH (1995)

Janet Heinrich, DrPH, MPH (1993)

Margaret H. Jordan, RN, MPH (1992)

C. William Keck, MD, MPH, *Immediate Past President*

Margo D. Kerrigan, MPH (1995)

Joyce C. Lashof, MD, *President*

Barry S. Levy, MD, MPH (1993)

Helen Rodriguez-Trias, MD, *President-Elect*

Allan G. Rosenfield, MD (1994)

Reed V. Tuckson, MD (1994)

Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO

William H. McBeath, MD, MPH, *Executive Director*

Joanne E. Lukomnik, MD, *Chair, Action Board*

George Pickett, MD, MPH, *Chair, Program Development Board*

George E. Hardy, MD, MPH. *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc., 1015 Fifteenth St., NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50,000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health,* 1015 Fifteenth St., NW, Washington, DC 20005.

Information on Manuscript Preparation is published in this Journal in January. The Journal publishes reports of original research, demonstrations, evaluations, and other articles covering current aspects of public health. Manuscripts are accepted for consideration with the understanding that they, or their essential substance, have not been published previously or accepted for publication and are not under consideration elsewhere.

Subscription price, individual, is $80/year US, $120/year international; **institutional,** $160/year US, $240/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to American Journal of Public Health, 1015 Fifteenth St., NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact: Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact: Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols. 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS Bibliographic Retrieval Service Inc, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc., Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Barrier Contraceptives and Sexually Transmitted Diseases in Women: A Comparison of Female-Dependent Methods and Condoms

Michael J. Rosenberg, MD, MPH, Arthur J. Davidson, MD, MSPH, Jian-Hua Chen, MD, MSPH, Franklyn N. Judson, MD, and John M. Douglas, MD

ABSTRACT

Introduction. Most efforts at sexually transmitted disease (STD) protection center on condom use, but little is known about how condoms compare with other barrier methods, particularly those controlled by women.

Methods. To evaluate the effect of different barrier contraceptives on the prevalence of STDs and other vaginal infections, we retrospectively studied 5681 visits by women to an urban STD clinic.

Results. As compared with women using no contraceptive or with tubal ligations, women using the contraceptive sponge or diaphragm had at least 65% lower rates or infection with *Neisseria gonorrhoeae* and *Trichomonas vaginalis,* while condom users had 34% and 30% lower rates, respectively. For *Chlamydia trachomatis,* the reduction was 13% among sponge users, 72% among diaphragm users, and 3% among condom users, although these differences were not significant. When compared with women using condoms, women using female-dependent methods (sponge or diaphragm) had significantly lower rates of both gonorrhea and trichomoniasis. Vaginal candidiasis was more common among women using diaphragms but not other barrier methods, while rates of bacterial vaginosis were similar among all groups.

Conclusions. Women using the contraceptive sponge or diaphragm experience protection from STDs to a greater extent than those relying on condoms. Female-dependent barrier contraceptives should receive more attention in STD risk-reduction programs. (*Am J Public Health.* 1992; 82:669–674)

Introduction

In 1988, an estimated 12 million new cases of sexually transmitted diseases (STDs) occurred in the United States. Three quarters of these were accounted for by three diseases—chlamydia (4 million cases), trichomoniasis (3 million), and gonorrhea (1.8 million).[1] The realizations that most STDs have increased over the past decade, that complications of these infections cost billions of dollars each year, and that certain STDs are risk factors for transmission of human immunodeficiency virus (HIV) have spurred efforts to identify effective means of preventing their spread. Such efforts are particularly important for women, since they suffer the greatest consequences of bacterial STDs in the form of pelvic inflammatory disease, ectopic pregnancy, and infertility.

Women who use barrier contraceptives (spermicides, the contraceptive sponge, condoms, or diaphragms) have generally lower prevalence rates of STDs than women using other forms of contraception or no contraception.[2] Clinical studies, however, show variable degrees of reduction in gonorrhea, nongonococcal urethritis, and chlamydia among users of condoms[3–8] and spermicides[6,7,9–14] as compared with contraceptive nonusers. This work is supported by laboratory evidence that intact condoms are impermeable to virtually all sexually transmitted organisms studied[15,16] and that spermicides effectively kill a variety of STD pathogens,[9] including HIV.[17]

Because the manner in which contraceptives are used in daily practice is an important determinant of protection against STDs, observational studies are an important reflection of protection applicable to large numbers of people. Two such studies indicated that female-dependent methods (spermicides and/or diaphragms) provided protection at least as strong as that afforded by condom use against gonorrhea.[6,18] These studies are limited, however, by absent or incomplete information on STD risk factors. In addition, most evaluated a single outcome, and all were performed before 1987, when prevention messages were not as widespread as they are today. To address these limitations, we studied the relationship between recent use of three popular forms of barrier contraceptives and prevalence of common STDs and vaginal infections among women attending a large urban STD clinic.

Methods

The Denver Metro Health Clinic for STDs serves 1.6 million residents of the Denver standard metropolitan statistical

Michael J. Rosenberg and Jian-Hua Chen are with Health Decisions Inc, Chapel Hill, NC. Michael J. Rosenberg is also with the Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill. Arthur J. Davidson, Franklyn N. Judson, and John M. Douglas are with the Disease Control Service, Denver Department of Public Health, Denver, Colo, and the Department of Preventive Medicine, University of Colorado Health Sciences Center, Denver. Franklyn N. Judson and John M. Douglas are also with the Department of Medicine, University of Colorado Health Sciences Center.

Requests for reprints should be sent to Michael J. Rosenberg, MD, MPH, Health Decisions, P.O. Box 4151, Chapel Hill, NC 27515.

This paper was submitted to the Journal February 27, 1991, and accepted with revisions October 2, 1991.

area. Women may visit the clinic because of genitourinary symptoms or concerns over STDs or may be referred by their partner, STD health investigators, another clinic, or the police department after an arrest for prostitution.

The study involved review of charts of women visiting the clinic for a new problem. Each woman was interviewed regarding demographic and behavioral information, including the single contraceptive that was used most frequently during the previous month. Routine laboratory evaluation of women included cervical culture for *Neisseria gonorrhoeae* and evaluation of vaginal fluid for pH and by microscopic examination. During certain periods, cervical culture for *Chlamydia trachomatis* was also performed as part of the routine examination; only visits by women during these periods were considered for the analysis of *C. trachomatis*.

Cultures for *N. gonorrhoeae* and *C. trachomatis* were carried out by conventional techniques.[19,20] The presence of *Trichomonas vaginalis* was detected by microscopic examination of saline suspension of vaginal secretions. A diagnosis of candidiasis was established by the presence of a characteristic curdy white vaginal discharge and/or the microscopic identification of fungal hyphae in vaginal discharge after treatment with potassium hydroxide. Bacterial vaginosis was diagnosed in women with symptoms of vaginitis by conventional diagnostic criteria[21] in the absence of another explanation for the symptoms.

After excluding women who used oral contraceptives (n = 1944), intrauterine devices (n = 173), or "other" (n = 134), or who had undergone a hysterectomy (n = 424), patients were divided into groups according to contraceptive used. For each outcome, women using one of three types of barrier contraceptives (sponge, diaphragm, or condom) were compared with a combined nonbarrier control group of women with a tubal ligation or using no contraception. Ninety percent test-based confidence limits[22] were used for the outcomes of gonorrhea, chlamydia, and trichomoniasis, consistent with the suggestion of protection from previous work.[3-14] Ninety-five percent confidence intervals (CIs) were used for outcomes of candidiasis and bacterial vaginosis, about both of which there is little prior information. CIs that exclude 1.0 are significant at the level of $P < .05$ (one-sided P values for 90% CIs, two-sided P values for 95% CIs).

Stepwise logistic regression analysis[23] was used to identify confounders[24] and to test for interaction.[25] This method retains those factors that change the odds ratio by 10% or more, although some variables were retained because of their clinical importance.

Results

Between January 1, 1987, and December 31, 1988, there were 5681 visits for the evaluation of a new problem by women using the contraceptive sponge, diaphragm, condom, tubal ligation, or no contraceptive. These visits occurred among 4162 women, of whom 73% had one visit during the study period, 16% had two visits, 6% had three visits, and 4% had four or more visits. *C. trachomatis* was the most prevalent infection (13.6%), followed by *Candida* (12.2%), *N. gonorrhoeae* (11.9%), *T. vaginalis* (11.1%), and bacterial vaginosis (10.7%). The majority (76%) of visits were by women using no barrier (tubal ligation or no contraceptive); condom use was reported at 18% of visits, diaphragm use at 4%, and the contraceptive sponge at 2% (Table 1). The women were predominantly young, heterosexual, and sexually active.

The most marked reductions were found among sponge users and diaphragm users for both gonorrhea (relative risk [RR] = 0.22 and 0.19, respectively) and trichomoniasis (RR = 0.23 and 0.20) (Table 2). A lesser but also significant degree of reduction was found among condom users for gonorrhea (RR = 0.70) and trichomoniasis (RR = 0.86). While both sponge and diaphragm users had a reduced prevalence of chlamydia (RR = 0.58 and 0.18), neither reduction was statistically significant given the fewer individuals evaluated. Condom users, in contrast, had no reduction in the prevalence of chlamydia (RR = 0.99). Adjustment for multiple risk factors altered the results little, slightly reducing estimates of protection for sponge and diaphragm users and minimally increasing those for condom users. Adjustment had the most marked effect when comparisons involved the fewest subjects (sponge users) and the least effect in the group with the greatest number of subjects (condom users), particularly for outcomes that were infrequently assessed (i.e., chlamydia).

For candidiasis and bacterial vaginosis, the only significant difference was found among diaphragm users, who had more candidal infections than did the control group (Table 2). Risk estimates for other contraceptive groups were close to 1.0, with 95% CIs that encompassed 1.0.

Because of the similarity of findings among women using the sponge and diaphragm, we combined these groups to compare female-dependent barrier methods with condoms. Women using female-dependent methods had significantly fewer infections with *N. gonorrhoeae* and *T. vaginalis* than did women who relied on their partners' use of condoms (Table 3). The combined group also had lower rates of *C. trachomatis* and bacterial vaginosis, although neither difference was statistically significant. Sponge and diaphragm users had a significantly higher prevalence of candidiasis than did condom users.

The minimal difference between unadjusted (crude) and adjusted risk estimates for the condom group indicates a lack of substantial confounding by factors included in the model. Where adjusting made the greatest difference, the change may reflect an unstable odds ratio rather than the adjusted estimate being a more accurate indicator. In addition, there were no meaningful changes in risk estimates when subgroups of women were analyzed according to whether women with at least one new partner in the past month were excluded, or whether only first visits during the 2-year study interval were considered (data not shown). We found no evidence of interaction between factors in the model.

Among all risk factors considered, contraceptives were the most consistent predictor of presence or absence of the STD outcomes studied (Table 4). Other risk factors studied had variable effects, depending on the outcome. For example, contact with a known or suspected STD carrier was a strong risk factor for gonorrhea, but not for trichomoniasis or chlamydia. The high number of significant associations for gonorrhea may, in part, reflect the greater number of subjects included in these analyses.

Discussion

Women using the contraceptive sponge, diaphragm, and condom all had significantly lower rates of infection with *N. gonorrhoeae* and *T. vaginalis* than did a control group of women who used no contraceptive or with tubal ligations. Each barrier contraceptive group also had lower rates of *C. trachomatis* infection, although the smaller number of cases limited the precision in defining this relationship. Furthermore, for gonorrhea and trichomoniasis, a greater reduction was found among

women using the contraceptive sponge or diaphragm than those using condoms. For the outcome of candidiasis, which might be considered an adverse consequence of use of spermicide-containing barrier contraceptives,[10] there was a significant elevation among diaphragm but not sponge users.

Our finding of lower disease rates among women using barrier methods is in agreement with previous literature. A study involving chart review of health maintenance organization patients found users of vaginal spermicides to have an 87% reduction in gonorrhea,[7] while a cross-sectional evaluation at an STD clinic found a 61% reduction.[12] Clinical trials of the contraceptive sponge[10,13] and of a vaginal spermicidal jelly[11] have reported reductions of 10% to 69% for gonorrhea and 21% to 40% for chlamydia.

Our work also indicates that the two female-dependent barrier methods investigated, the contraceptive sponge and diaphragm, may provide better user effectiveness against various STDs than condoms. Such a finding was suggested by an earlier case-control study that found a 13% reduction in gonorrhea among condom users, a 10% reduction among users of spermicides alone, and a 55% reduction among women who used diaphragms in conjunction with spermicides.[6] Other studies found significant reductions of 64% for gonorrhea and 97% for chlamydia among diaphragm users, but reductions among condom users were reported only as nonsignificant.[18] Trichomoniasis was found to be 80% lower in spermicidal foam users but only 25% and 22% lower in diaphragm and condoms users, respectively.[26] We found a significantly lower prevalence of gonorrhea and trichomoniasis among users of both diaphragms (which we assume were used with spermicides) and sponges than among users of condoms that persisted after adjustment for multiple STD risk indicators.

There is less information regarding the relationship of barrier contraceptives and candidiasis, trichomoniasis, and bacterial vaginosis. Candidiasis was found nearly three times more frequently among contraceptive sponge users than in women not using a vaginal contraceptive in a clinical trial conducted among prostitutes in Thailand.[10] These women used the sponge nearly continuously until they either became infected with gonorrhea or chlamydia or completed a 3-week period. In contrast, most sexually active women are likely to use barrier contraceptives only intermittently, perhaps allowing a greater opportunity for normal vaginal bacterial flora to reestablish colonization. This possibility is suggested by our finding of an increased rate of candidiasis among diaphragm but not sponge users, as well as other studies that found minimal or no change in candidiasis among spermicide users.[25,27,28]

The few studies of bacterial vaginosis indicate modest, nonsignificant reductions among spermicide users.[21,28] Our study, which is substantially larger, indicates a virtually identical rate of bacterial vaginosis among users of sponges, diaphragms, and condoms and among the control group, suggesting that none of these barrier methods are effective in preventing this syndrome and that spermicides have little effect on the alteration in vaginal flora associated with bacterial vag-

TABLE 1—Characteristics and Sexual Habits by Contraceptive Used

Characteristic	Sponge (n = 101)	Diaphragm (n = 227)	Condom (n = 1021)	None or Tubal Ligation (n = 4332)
Age, mean y	27.5	29.1	23.9	26.5
Reason for visit, %				
Evaluation of symptoms	61	63	46	49
Contact of person with sexually transmitted disease	27	21	27	35
Asymptomatic	7	12	7	6
Treatment for diagnosis made elsewhere	5	3	4	6
Vice squad referral	0	1	16	4
Ethnic group, %				
White	72	75	57	42
Black	15	14	29	33
Hispanic	11	10	13	23
Other	2	1	1	2
New partners in past month, %				
0	71	69	66	76
1	25	26	16	18
2–5	3	4	7	5
6+	1	1	11	1
Partners in previous month, %				
0	9	9	9	12
1	69	70	55	66
2–5	21	20	23	21
6+	1	1	13	1
Lifetime partners, %				
0	1	0	0	0
1–3	8	7	15	17
4–6	11	15	19	23
7–9	9	6	7	8
10–20	42	38	20	25
21+	13	23	23	12
Declined to answer	16	11	16	15
Sexual activity in past month, %				
Genital sex	99	100	99	99
Oral sex, active	43	48	42	31
Oral sex, passive	43	47	41	31
Anal sex	5	3	2	2
Sexual preference, %				
Heterosexual	100	99	98	97
Bisexual	0	0	1	1
Homosexual	0	1	1	2
Previous sexually transmitted diseases, %				
Gonorrhea	22	24	41	39
Trichomoniasis	22	31	34	32
Crabs	13	21	20	15
Genital warts	12	15	13	9
Chlamydia	10	11	13	8
Pelvic inflammatory disease	11	10	13	10
Bacterial vaginosis	8	14	11	8
Genital herpes	7	11	5	6
Hepatitis	2	3	3	2
Scabies	2	1	2	2
Syphilis	0	0	2	2

TABLE 2—Risk of Clinical Outcomes by Type of Barrier Contraceptive

Outcome	Sponge (n = 101)[a] % Infected	Sponge (n = 101)[a] Odds Ratio Crude	Sponge (n = 101)[a] Odds Ratio Adjusted[e]	Sponge (n = 101)[a] Confidence Interval[f]	Diaphragm (n = 227)[b] % Infected	Diaphragm (n = 227)[b] Odds Ratio Crude	Diaphragm (n = 227)[b] Odds Ratio Adjusted[e]	Diaphragm (n = 227)[b] Confidence Interval[f]
Gonorrhea	3.1	0.22	0.31	0.11–0.81*	2.7	0.19	0.32	0.16–0.65*
Trichomoniasis	3.0	0.23	0.29	0.11–0.76*	2.6	0.20	0.24	0.12–0.48*
Chlamydia	8.7	0.58	0.64	0.19–2.24	2.9	0.18	0.25	0.05–1.36
Any of above	7.9	0.28	0.37	0.20–0.69*	5.3	0.18	0.25	0.15–0.42*
Bacterial vaginosis	10.9	1.06	1.15	0.60–2.18	12.3	1.22	1.16	0.76–1.75
Candidiasis	13.9	1.20	1.20	0.68–2.14	19.8	1.84	1.88	1.33–2.65*

Outcome	Condom (n = 1021)[c] % Infected	Condom (n = 1021)[c] Odds Ratio Crude	Condom (n = 1021)[c] Odds Ratio Adjusted[e]	Condom (n = 1021)[c] Confidence Interval[f]	None+Tubal Ligation (n = 4332)[d] % Infected	None+Tubal Ligation (n = 4332)[d] Odds Ratio Crude	None+Tubal Ligation (n = 4332)[d] Odds Ratio Adjusted[e]
Gonorrhea	9.4	0.70	0.61	0.50–0.76*	12.8	Referent	
Trichomoniasis	10.4	0.86	0.80	0.65–0.97*	11.9	Referent	
Chlamydia	14.0	0.99	0.93	0.60–1.44	14.1	Referent	
Any of above	18.8	0.76	0.68	0.59–0.80*	23.4	Referent	
Bacterial vaginosis	11.6	1.13	1.21	0.97–1.53	10.4	Referent	
Candidiasis	12.1	1.02	0.99	0.79–1.23	11.8	Referent	

Note. Confidence intervals are based on adjusted relative risks.
[a]Except n = 97 for gonorrhea, n = 23 for chlamydia.
[b]Except n = 225 for gonorrhea, n = 35 for chlamydia.
[c]Except n = 984 for gonorrhea, n = 157 for chlamydia.
[d]Except n = 4068 for gonorrhea, n = 4326 for trichomoniasis, n = 765 for chlamydia, n = 4330 for any of above and for candidiasis.
[e]Adjusted for age, race, number of sexual partners in previous month, and history of past trichomoniasis, chlamydia, or gonorrhea.
[f]90% confidence interval for gonorrhea, trichomoniasis, chlamydia, and any of above; 95% confidence interval for bacterial vaginosis and candidiasis.
*$P < .05$.

TABLE 3—Risk of Clinical Outcome for Users of Contraceptive Sponge or Diaphragm as Compared with Users of Condoms

Outcome	Odds Ratio Crude	Odds Ratio Adjusted[a]	95% Confidence Interval
Gonorrhea	0.28	0.45	0.22–0.92*
Trichomoniasis	0.24	0.33	0.17–0.64*
Chlamydia	0.33	0.84	0.25–2.81
Bacterial vaginosis	1.03	0.88	0.59–1.31
Candidiasis	1.60	1.62	1.15–2.28*

[a]Adjusted for age, race, number of sexual partners in previous month, number of new partners in previous month, total lifetime partners, symptoms, reason for visit, and history of chlamydia, gonorrhea, or trichomoniasis.
*$P(2) < .05$.

inosis. Our finding of a reduced rate of trichomoniasis is similar to that from a cross-sectional study[29] and less marked than the modest, nonsignificant decrease for participants in a placebo-controlled clinical trial.[28]

The size and design of this study provide several advantages over previous efforts. First, our data reflect day-to-day, subject-reported use in a nonmonitored environment. Such user effectiveness, as differentiated from efficacy in optimal circumstances (method effectiveness), is an important consideration in judging usefulness in STD prevention efforts. Second, the large number of women and three different STD outcomes analyzed allows a more comprehensive evaluation of the effect of individual contraceptives, including a recent introduction, the contraceptive sponge, on a range of outcomes. Finally, the detailed information on STD risk indicators provides assurance that the observed relationships are not distorted by STD risk that may have differed between groups of women using different contraceptives.

Although we had extensive information on factors related to risk of STD acquisition, there may be other unquantified factors related to STD risk or contraceptive choice for which we were unable to control. For example, we do not know how reliably the STD risk factors and markers we measured characterized risk profiles of women or their partners and whether unmeasured factors might have systematically differed between contraceptive user groups. However, the consistency of factors associated with STDs in this study (age, race, reason for visit, number of new sexual partners, and total number of sexual partners) and other studies and the relatively small change in risk estimates after adjusting for these factors suggest a lack of substantial bias.

Misclassification error may have been introduced by the lack of information on concurrent use of more than one contraceptive method or inconsistent use of contraceptives. Also, we evaluated only a relatively short time period (1 month), and women may have become infected prior to this time. This may be particularly important for an infection such as *C. trachomatis,* which may have a prolonged asymptomatic carrier state. If such errors

were common, they could cause an over- or underestimate of true relative risk, depending on the sequence, the effectiveness of each method, and how each contraceptive was used. We suspect that simultaneous use of more than one method is probably rare, based on a previous study of contraception and gonorrhea that found that only 2% of women used some form of spermicides in addition to condoms.[6] If women more frequently used multiple methods, and the use of multiple methods were approximately equal in different user groups, then the differences reflected in our study would underestimate the true magnitude of protection. Similarly, inconsistent use of a single method would also underestimate the reduction of risk provided by barrier contraceptives, with the magnitude proportional to the degree of nonuse. Of interest, however, is the finding during a recent clinical trial of spermicides that only slight reduction of risk occurred with increasing compliance.[11] This observation suggests that protective measures are more likely to be used during high-risk encounters and is supported by our finding of no change in risk after women with a new partner in the previous month were eliminated from the analysis.

An additional source of bias due to disease ascertainment could have arisen from the fact that clinicians were aware of the contraceptive being used. However, since diagnostic studies performed were not influenced by the contraceptive used, this factor is unlikely to have influenced the results. Finally, it is also possible that women using different contraceptive methods may be more likely to seek STD care at other health care facilities. If such care were more common among women using the contraceptive sponge or diaphragm, our findings would overestimate the associated degree of reduction.

The difference in STD rates between female- and male-dependent barriers may be partially due to different frequencies of use. The disparity between the high levels of condom efficacy in the laboratory[15] and the results of this and other studies suggests that condoms are often used improperly and/or intermittently. This finding is consistent with reports of low condom use, even in high-risk groups.[30-35] A recent review demonstrates underuse of condoms, particularly among individuals at high risk such as persons with multiple sex partners and those previously treated for an STD.[36]

Our data suggest that efforts to help control STDs by use of barrier contraceptives should be oriented toward women to a greater extent than in the past. STDs can be transmitted more easily from men to women than vice versa,[5,37,38] and women experience more frequent and serious complications. Although STD prevention efforts have emphasized condoms to the virtual exclusion of other barrier methods, there is a need to identify methods that women can use independently of their partners. This is particularly true for those at highest risk for STDs, since these women, often young or socioeconomically disadvantaged, may lack the interpersonal skills necessary to negotiate the use of condoms. Identifying methods that are available without prescription is especially important for adolescents, who are at highest risk of pregnancy and STDs. Newer spermicidal preparations, such as the contraceptive sponge, may help STD prevention efforts because of their better acceptability as compared with other vaginal preparations (A. C. Nielsen, Drug and Food Index data, unpublished), and thus their greater likelihood of use.

The urgency of the STD problem and the incomplete information about the value of female-dependent methods[39,40] indicate the importance of further investigations to confirm and extend our findings. Most immediately, studies are needed that will provide more complete information about contraceptive use and STD risk indicators of women and their sexual partners. While prospective studies will generate useful data, it must be recognized that the intensive intervention that occurs in such studies will result in contraceptive use patterns that are unlikely to be achieved in everyday use. Thus, observational studies, with emphasis on careful validation of STD risk and contraceptive use, will remain a vital component of such needed research. □

TABLE 4—Odds Ratios for Factors Related to Clinical Outcomes[a]

	Gonorrhea (n = 5374)	Trichomoniasis (n = 5675)	Chlamydia (n = 980)	Bacterial Vaginosis (n = 5681)	Candidiasis (n = 5679)
Sponge use	0.29*	0.26*	0.87	0.94	1.15
Diaphragm use	0.35*	0.29*	0.28	0.99	1.82*
White race[b]	0.41*	0.82	0.59*	0.84	1.25
Vice squad referral	0.49*	2.68*	1.82	0.61	0.47*
Condom use	0.66*	0.70*	0.97	1.11	0.95
Presence of genitourinary symptoms	0.90	1.50*	1.49	2.38*	2.53*
Age[c,d]	0.93*	1.01	0.93*	1.03*	0.98
History of gonorrhea, trichomoniasis, or chlamydia	0.97	1.32*	0.63*	1.44*	0.89
No. of new partners, past month[e]	1.15*	1.00	1.16	0.85*	0.94
No. of partners, past month[f]	1.22*	1.06	0.95	1.17*	1.06
Black race[g]	1.84*	2.16*	0.72	1.01	1.34*
Contact with known or suspected STD carrier	3.61*	1.07	1.24	0.33*	0.59*

Note. STD = sexually transmitted disease.
[a]Risk of outcome, adjusted by each of the other factors listed.
[b]Compared with Hispanic, Asian, Native American, Black, and other non-White.
[c]Change for each successive year.
[d]For all visits, mean = 25.8 years (SD = 7.5).
[e]For all visits, mean = 1.3 (SD = 1.0).
[f]For all visits, mean = 0.4 (SD = 1.0).
[g]Compared with Hispanic, Asian, Native American, White, and other non-Black.
*$P(2) < .05$, except sponge, diaphragm, and condom for outcomes of gonorrhea, trichomoniasis, and chlamydia, where $P(1) < .05$.

Acknowledgments

This work was funded in part by a grant from Wyeth International Limited.

Portions of this paper were presented at the meeting of the International Society for Sexually Transmitted Disease Research, Copenhagen, Denmark, September 1989.

We are grateful for the statistical assistance of Greg Samsa, PhD, the efforts of Carol Root, RN, Nurse Supervisor, and the clinical and laboratory staff of the Denver Metro Health Clinic; and the thoughtful manuscript review by Willard Cates, Jr, MD, MPH.

References

1. Division of Sexually Transmitted Diseases. *Annual Report, 1989.* Atlanta, Ga: Centers for Disease Control; 1990.
2. Stone KM, Grimes D, Magder LS. Personal protection against sexually transmitted diseases. *Am J Obstet Gynecol.* 1986;155:180–188.
3. Pemberton J, McCann JS, Mahoney JDH, MacKenzie G, Dougan H, Hay I. Sociomedical characteristics of patients attending a VD clinic and the circumstances of infection. *Br J Vener Dis.* 1972;48:391–396.
4. Barlow D. The condom and gonorrhea. *Lancet.* 1977;2:811–812.
5. Hooper RR, Reynolds GH, Jones OG, et al. Cohort study of venereal disease. I: The risk of gonorrhea transmission from infected women to men. *Am J Epidemiol.* 1978;108:136–143.
6. Austin H, Louv WC, Alexander J. A case-control study of spermicides and gonorrhea. *JAMA.* 1984;251:2822–2824.
7. Jick H, Hannan MT, Stergachis A, Heidrich F, Perera DR, Rothman KJ. Vaginal spermicides and gonorrhea. *JAMA.* 1982;248:1619–1621.
8. Darrow WW. Condom use and use-effectiveness in high-risk populations. *Sex Transm Dis.* 1989;16:157–160.
9. Rosenberg MJ, Feldblum PF. Do spermicides protect against sexually transmitted diseases? *Afr J Sex Transm Dis.* 1986;2:42–45.
10. Rosenberg MJ, Rojanapithayakorn W, Feldblum PG, Higgins JE. Effect of the contraceptive sponge on chlamydial infection, gonorrhea, and candidiasis. *JAMA.* 1987;257:2308–2312.
11. Louv WC, Austin H, Alexander WJ, Stagno S, Cheeks J. A clinical trial of nonoxynol-9 for preventing gonococcal and chlamydia infections. *J Infect Dis.* 1988; 158:518–523.
12. Quinn RW, O'Reilly KR. Contraceptive practices of women attending the sexually transmitted disease clinic in Nashville, Tennessee. *Sex Transm Dis.* 1985;12:99–102.
13. Rosenberg MJ, Feldblum PH, Rojanapithayakorn W, Sawasdivorn W. The contraceptive sponge's protection against *Chlamydia trachomatis* and *Neisseria gonorrhoeae*. *Sex Transm Dis.* 1987; 14:147–152.
14. Magder LS, Harrison HR, Ehret JM, Anderson TS, Judson FN. Factors related to genital *Chlamydia trachomatis* and its diagnosis by culture in a sexually transmitted disease clinic. *Am J Epidemiol.* 1988; 28:298–308.
15. Centers for Disease Control. Condoms for prevention of sexually transmitted diseases. *MMWR.* 1988;37:133–137.
16. Rietmeijer CAM, Krebs MS, Feorina PM, et al. Condoms as physical and chemical barriers against HIV. *JAMA.* 1988; 259:1851–1853.
17. Hicks DR, Martin JP, Getchell JP, et al. Inactivation of HTLV-III/LAV-infected cultures of normal human lymphocytes by nonoxynol-9 in vitro. *Lancet.* 1985;2:1422–1423.
18. Bradbeer CS, Thin RN, Tan R, Thirumoorthy T. Prophylaxis against infection in Singaporean women. *Genitourin Med.* 1988;64:52–53.
19. Ehret JM, Knapp JS. Gonorrhea. *Clin Lab Med.* 1989;9:445–480.
20. Ehret JM, Judson FN. Genital *Chlamydia* infections. *Clin Lab Med.* 1989;9:481–500.
21. Amsel R, Totten PA, Spiegel CA, Chen KCS, Eschenbach D, Holmes KK. Nonspecific vaginitis: diagnostic criteria and microbial and epidemiologic associations. *Am J Med.* 1983;74:14–22.
22. SAS Institute Inc. *SAS Users Guide.* Cary, NC: SAS Institute Inc; 1988.
23. Harrell FE. The LOGIST procedure. In: *SUGI Supplemental Library User's Guide, Version 5 Edition.* Cary, NC: SAS Institute Inc; 1986:125–137.
24. Mickey RM, Greenland S. A study of the impact of confounder-selection criteria on effect estimation. *Am J Epidemiol.* 1989;129:125–137.
25. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiological research: principles and quantitative methods. New York, NY: Van Nostrand Reinhold; 1982:403–418.
26. Lazar A. *Trichomonas vaginalis* infection: incidence with use of various contraceptive methods. *J Med Soc NJ.* 1970;67:225–226.
27. Foxman B. The epidemiology of vulvovaginal candidiasis: risk factors. *Am J Public Health.* 1990;80:329–331.
28. Barbone F, Austin H, Louv WC, Alexander WJ. A follow-up study of methods of contraception, sexual activity, and rates of trichomoniasis, candidiasis, and bacterial vaginosis. *Am J Obstet Gynecol.* 1990; 163:510–514.
29. Feldblum PJ, Bernardik E, Rosenberg MJ. Spermicide use and sexually transmitted disease. *JAMA.* 1988;259:2851.
30. Lawrence DN, Jason JM, Holman RC, Heine P, Evatt BL. Sex practice correlates of human immunodeficiency virus transmission and acquired immunodeficiency syndrome incidence in heterosexual partners and offspring of U.S. hemophiliac men. *Am J Hematol.* 1989;30:68–76.
31. Fischl MA, Dickinson GM, Scott GB, et al. Evaluation of heterosexual partners, children, and household contacts of adults with AIDS. *JAMA.* 1987;257:640–644.
32. Quinn TC, Glasser D, Cannon RO, et al. Human immunodeficiency virus infection among patients attending clinics for sexually transmitted diseases. *N Engl J Med.* 1987;318:197–203.
33. Mann J, Quinn TC, Piot P, et al. Condom use and HIV infection among prostitutes in Zaire. *N Engl J Med.* 1987;316:345.
34. Centers for Disease Control. Positive HTLV-III/LAV antibody results for sexually active female members of social/sexual clubs. *MMWR.* 1986;35:697–699.
35. Conant MA. Condoms and the prevention of AIDS. *JAMA.* 1986;256:1442–1443.
36. Solomon MZ, DeJong W. Preventing AIDS and STDs through condom promotion: a patient education intervention. *Am J Public Health.* 1989;79:453–458.
37. Holmes KK, Johnson DW, Trostle HJ. An estimate of the risk of men acquiring gonorrhea by sexual contact with infected females. *Am J Epidemiol.* 1970;91:17–24.
38. Platt R, Rice PA, McCormack WM. Risk of acquiring gonorrhea and prevalence of abnormal adnexal findings among women recently exposed to gonorrhea. *JAMA.* 1983;250:3205–3209.
39. Stein ZA. HIV prevention: the need for methods women can use. *Am J Public Health.* 1990;80:460–462.
40. National Research Council, National Academy of Sciences. *AIDS: The Second Decade.* Washington, DC: National Academy Press; 1990:112–113.

Pregnancy-Induced Hypertension in North Carolina, 1988 and 1989

David A. Savitz, PhD, and Jun Zhang, MB

ABSTRACT

Introduction. Pregnancy-induced hypertension (PIH) is a highly prevalent pregnancy complication with adverse effects on maternal and infant health. Epidemiologic research concerning its etiology is limited.

Methods. Birth records from North Carolina for the period 1988 through 1989 include an indication of the presence of PIH. The risk of PIH was examined in relation to several maternal characteristics and exposures, including reproductive history, demographic characteristics, and tobacco use during pregnancy. Risk ratio estimates, adjusted for confounders, were calculated contrasting PIH among exposed vs unexposed women.

Results. The overall risk of PIH was 43.1 per 1000 births, with multiple pregnancies, nulliparity, and advanced maternal age associated with markedly increased risks. Tobacco use was inversely associated with PIH, and Blacks and Whites were at virtually equal risk.

Conclusions. Problems in diagnosis and classification impede research in this area, with birth certificates limited in quality and breadth of information. Nonetheless, several patterns emerged that are worthy of further epidemiologic evaluation using more sophisticated designs. (*Am J Public Health* 1992;82:675–679)

Introduction

Pregnancy-induced hypertension (PIH) is among the most common pregnancy complications, with a cumulative incidence estimated as high as 7%.[1,2] PIH is associated with a substantial increase in growth retardation and preterm birth and threatens the life of the mother if untreated.[3] Classification of PIH remains controversial, but most authorities recognize that (1) preexisting hypertension and hypertension developing during labor are distinct from hypertension developing during pregnancy[4] and (2) preeclampsia, in which the blood pressure elevation is accompanied by proteinuria or edema, is distinct from gestational hypertension alone.[5] Nonetheless, the common occurrence of edema in late pregnancy and frequent difficulty in accurately separating preeclampsia from isolated gestational hypertension[4] make PIH a suitable entity for etiologic research, while recognizing that there are likely to be heterogeneous subcategories present.[6]

There has been limited research into the etiology of PIH, in spite of its frequency. Primiparas are known to be at markedly greater risk of preeclampsia than multiparas.[7] Women who smoke cigarettes have been reported to show a reduced risk of preeclampsia.[8,9] Elevated blood lead levels,[10] exposure to organic solvents,[11] lack of leisure-time physical activity early in pregnancy,[12] male fetuses,[13] low levels of calcium intake,[14] use of barrier contraceptives,[15] and young maternal age[16] have all been reported to increase the risk of PIH or preeclampsia, but these observations await corroboration or refutation. These clues deserve follow-up in focused studies, yet additional analyses that simply describe the patterns of occurrence in defined populations are also needed. Even such basic questions as the presence of racial differences in risk have not been resolved.[16] The modification of the North Carolina birth certificate in 1988 to include a checklist item for PIH, as well as several risk factors of potential interest (e.g., cigarette smoking), provides an efficient opportunity to examine patterns of PIH among women in the state.

Methods

Birth data for North Carolina for the years 1988 and 1989 were made available from the Center for Health and Environmental Statistics of the North Carolina Department of Environment, Health, and Natural Resources. A nurse or clerk in the hospital of delivery completes a certificate of live birth, which includes the usual array of demographic and residential information as well as a series of health conditions in check-box format. Three items from the birth certificate were used in this analysis: "hypertension, chronic"; "hypertension, pregnancy-associated"; and "eclampsia." This information is obtained from the medical record subsequent to delivery, with no formal quality-control procedures. Nonetheless, the clinical significance of these events makes them likely to be noted with some accuracy and prominence in the charts. Additional items used in the analysis were mother's age, moth-

The authors are with the Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill and the Carolina Population Center.

Requests for reprints should be sent to David A. Savitz, PhD, Department of Epidemiology, CB #7400, School of Public Health, University of North Carolina, Chapel Hill, NC 27599-7400.

This paper was submitted to the Journal September 30, 1991, and accepted with revisions December 10, 1991.

er's education (in years), child's race (Black, White, other), child's sex, gravidity, parity, plurality (singleton, multiple birth), onset of prenatal care, tobacco use (cigarettes per day), gestational age, birthweight, weight gain during pregnancy, mother's residence (urban/rural, Coastal Plain vs other), and month and season of birth. Although information on alcohol use is requested, the very small number of women who were reported to have used alcohol during pregnancy (2.6%) suggests that the information is unsuitable for analysis.

Starting with a total of 199 651 live births to residents of North Carolina, women with missing data on PIH (n = 1636), preexisting hypertension present (n = 1869), or missing data on preexisting hypertension (n = 1636) were omitted (some women having more than one exclusion criterion), leaving 196 146 births for analysis. Risk among women with multiple as compared with singleton gestations differed markedly (80.6 per 1000 vs 42.2 per 1000, respectively), so women with multiple fetuses (n = 4527) were omitted from the main analyses. Absence of prenatal care was associated with reduced incidence of PIH (risks of 43.4 per 1000 and 22.4 per 1000, respectively). This pattern is presumably due to lack of opportunity for diagnosis, so that women with no prenatal care or with missing data on the onset of prenatal care were also omitted (n = 3975). This left 187 742 women (91 188 in 1988 and 96 554 in 1989) with information on PIH, known absence of preexisting hypertension, singleton gestations, and known initiation of prenatal care prior to delivery. PIH was defined as an indication of either "hypertension, pregnancy-associated" or "eclampsia." By definition, eclampsia must be preceded by preeclampsia.[4,6]

Risks of PIH were calculated per 1000 live births by dividing number of PIH cases by the total number of births in the stratum. We examined risk by calendar year, season, and geographic area in the state (focusing on urban–rural gradients and residence in the Coastal Plain). Demographic characteristics of interest included mother's race (White, Black, other), mother's and father's age, and mother's education. Races other than Black and White were too diverse for analysis, with small numbers of American Indians (2.5% of births), Filipinos (0.5% of births), and other Asian or Pacific Islanders (0.3% of births) and a smaller number of Chinese, Japanese, and Hawaiians.

Gravidity and parity were examined as indicators of reproductive history. Attributes of the present pregnancies included infant sex, timing of prenatal care onset, smoking during pregnancy, and potential consequences of PIH, including weight gain during pregnancy, gestational age at delivery, and birthweight.

Crude risks and risk ratios were calculated and adjusted odds ratios (ORs) obtained by Mantel-Haenzel adjustment and logistic regression.[17] PIH is sufficiently rare in this population (under 5%) that risk ratios and ORs would be similar in magnitude.

Results

For the period 1988 through 1989, the risk of PIH in North Carolina was 43.1 per thousand. Although lower than is often reported, this highly prevalent problem affected over 4200 women in the state each year. Risks were similar over the 2 years, with women delivering in the summer or fall at somewhat lower risk than those delivering in the winter or spring (40.3 and 46.1 per thousand, respectively).

Table 1 presents crude risk ratios and risk ratios adjusted for parity, with 95% confidence intervals (CIs). Dividing the state's population by community size yielded indications of a slightly higher risk in the most rural areas. Comparison of the Coastal Plain with the remainder of the state (data not shown) identified no gradient in PIH risk (adjusted OR = 1.0, 95% CI = 1.0–1.1).

A monotonically increasing risk of PIH was observed in relation to maternal age, with a twofold difference between the youngest (<20) and oldest (40+) mothers. Little difference was found across levels of maternal education, with a slight decrease in risk for women with 16 or more years of education. Blacks and Whites had similar PIH risks, with other races slightly lower. Infant sex was unrelated to PIH.

Gravidity and parity showed the most pronounced patterns: first pregnancies and births had double the risk of PIH compared with subsequent ones, with no evidence of a dose–response gradient after the first. Slight indications of a rising risk beyond parity 3 were observed. Among nulliparous women, there was no effect of differing gravidity: there was a relative risk of 1.0 for nulliparous women with previous pregnancies compared with those without previous pregnancies. Among mothers who received prenatal care, there were no differences in relation to the trimester in which care was begun.

Crude results for tobacco use showed an inverse gradient with PIH risk, which was diminished by adjustment for parity. Even after adjustment, however, there was a monotonically decreasing risk, with increasing amount smoked culminating in an adjusted risk ratio of 0.7 in women smoking 20 or more cigarettes per day.

Pronounced associations were found between PIH and reduced gestational age, reduced birthweight, and increased weight gain during pregnancy. These are not thought to be pertinent to the etiology of PIH; rather, they are presumably a reflection of the disease itself (increased weight gain due to edema) or the consequences of PIH (gestational age, birthweight).

Variables that were found to be associated with PIH (parity, mother's age, smoking during pregnancy) and others of a priori interest (race, mother's education) were examined simultaneously in a logistic regression model (Table 2). The results were generally consistent with the parity-adjusted risk ratios. Nulliparity remained strongly associated with increased risk. Blacks and Whites were indistinguishable in risk, with other races at lower risk. Older maternal age was associated with increased PIH risk, and education above 16 years was associated with decreased risk. Prenatal care onset was still essentially unrelated to PIH. Smoking's association with reduced risk was not affected by adjustment.

Separate logistic regression analyses for Blacks and Whites yielded similar patterns, except that the increased risk with advanced maternal age was stronger among Blacks and the benefits of high levels of education were stronger among Whites (Table 2).

Results were also compared for women of parity 0 vs women of parity 1+ (Table 2). The increased risk of PIH with increased maternal age was notably weaker among nulliparous women but was still apparent. Otherwise, the pattern of risk factors among the two parity groups was quite similar.

Discussion

The overall risk of PIH found in this analysis of North Carolina birth certificates (43.1 per 1000 singleton live births) is slightly lower than has been reported in the literature[2] but compatible with recent estimates of 26 per 1000 births for preeclampsia.[16] There is substantial confusion due to inconsistent terminology and diagnostic inaccuracies among preexisting

TABLE 1—Crude and Adjusted Risk Ratios for PIH in Relation to Demographic Characteristics, Reproductive History, Behaviors during Pregnancy, and Pregnancy Outcome: North Carolina, 1988–1989

	No. of Cases	No. of Births	Risk	Crude Risk Ratio	Adjusted Risk Ratio[a]	95% Confidence Interval
Community population						
Rural	4670	102 162	45.7	1.0	1.0	...
2500–4999	433	9928	43.6	1.0	1.0	0.9–1.1
5000–9999	360	9621	37.4	0.8	0.8	0.7–0.9
10 000–24 999	626	14 585	42.9	0.9	0.9	0.9–1.0
25 000–49 999	341	10 404	32.8	0.7	0.7	0.6–0.8
50 000+	1563	41 042	38.1	0.8	0.8	0.8–0.9
Mother's age, y						
<20	1522	30 278	50.3	1.2	0.9	0.9–1.0
20–29	4655	113 053	41.2	1.0	1.0	...
30–34	1241	32 593	38.1	0.9	1.1	1.0–1.2
35–39	473	9946	47.6	1.2	1.5	1.3–1.6
40+	72	1281	56.2	1.4	1.9	1.5–2.4
Mother's education, y						
<9	257	6137	41.9	0.9	0.9	0.8–1.0
9–11	1486	36 498	40.7	0.9	0.9	0.8–0.9
12	3506	77 705	45.1	1.0	1.0	...
13–15	1614	35 823	45.1	1.0	1.0	0.9–1.1
16+	1122	31 327	35.8	0.8	0.8	0.7–0.8
Child's race						
White	5560	127 708	43.5	1.0	1.0	...
Black	2265	54 655	41.4	1.0	1.0	1.0–1.0
Other	168	5379	31.2	0.7	0.7	0.6–0.9
Child's sex						
Male	4153	95 816	43.3	1.0	1.0	...
Female	3840	91 926	41.8	1.0	1.0	0.9–1.0
Gravidity						
0	4383	71 185	61.6	1.0	...	..
1	1948	59 212	32.9	0.5	...	0.5–0.6[b]
2	915	32 548	28.1	0.5	...	0.4–0.5
3	405	14 445	28.0	0.5	...	0.4–0.5
4+	330	10 036	32.9	0.5	...	0.5–0.6
Parity						
0	5342	87 360	61.1	1.0	...	...
1	1716	63 474	27.0	0.4	...	0.4–0.5[b]
2	619	25 863	23.9	0.4	...	0.3–0.4
3	202	7622	26.5	0.4	...	0.4–0.5
4+	114	3374	33.8	0.6	...	0.5–0.7
Prenatal care onset, mo						
1–3	6328	144 395	43.8	1.0	1.0	...
4–6	1389	35 788	38.8	0.9	0.9	0.9–1.0
7–9	276	7569	36.5	0.8	0.9	0.8–1.0
Tobacco use during pregnancy, no. of cigarettes per day						
0	6604	148 345	44.5	1.0	1.0	...
1–9	359	9460	37.9	0.9	0.9	0.8–1.0
10–19	459	14 307	32.1	0.7	0.8	0.7–0.8
20+	369	12 950	28.5	0.6	0.7	0.7–0.8
Amount unknown	129	3496	36.9	0.8	0.9	0.7–1.0
Gestational age, wk						
<37	1125	19 693	57.1	1.4	1.4	1.3–1.4
37–41	5955	142 671	41.7	1.0	1.0	...
42+	913	25 378	36.0	0.9	0.8	0.8–0.9
Birthweight, g						
<1500	191	2152	88.8	2.4	2.3	2.0–2.6
1500–2500	828	9962	83.1	2.2	2.2	2.0–2.3
2501–3000	1397	30 476	45.8	1.2	1.2	1.1–1.3
3001–3500	2587	69 675	37.1	1.0	1.0	...
3501–4000	2121	55 190	38.4	1.0	1.1	1.0–1.1
4001+	868	20 168	43.0	1.2	1.2	1.1–1.3
Weight gain, lb						
0–20	1359	38 305	35.5	0.9	1.0	0.9–1.1
21–40	4213	110 180	38.2	1.0	1.0	...
41–60	1608	24 083	66.8	1.7	1.6	1.5–1.7
61+	278	2274	122.3	3.2	2.9	2.6–3.2

Note. PIH = pregnancy-induced hypertension. Subjects with missing data on PIH, preexisting hypertension, or prenatal care; preexisting hypertension present; multiple gestation; and no prenatal care were excluded.
[a]Adjusted for parity (1 vs 2+).
[b]Confidence intervals on crude risk ratios.

Table 2—Results of Logistic Regression of Pregnancy-Induced Hypertension by Race and Parity: North Carolina, 1988–1989

	Total		Whites		Blacks		Parity 0		Parity 1+	
	Adjusted Odds Ratio	95% Confidence Interval	Adjusted Odds Ratio	95% Confidence Interval	Adjusted Odds Ratio	95% Confidence Interval	Adjusted Odds Ratio	95% Confidence Interval	Adjusted Odds Ratio	95% Confidence Interval
Parity (1+ vs 0)	0.4	0.3–0.4	0.4	0.3–0.4	0.4	0.3–0.5	...	...	...	...
Race										
Black vs White	1.0	0.9–1.0	...	...	...	...	0.9	0.8–1.0	1.1	1.0–1.2
Other vs White	0.7	0.6–0.9	...	...	...	...	0.7	0.5–0.9	0.7	0.5–1.0
Mother's age, y										
<20 vs 20–34	0.9	0.8–0.9	0.9	0.8–1.0	0.9	0.7–1.0	1.0	0.8–1.0	0.5	0.4–0.7
35+ vs 20–34	1.6	1.4–1.8	1.4	1.2–1.6	2.1	1.7–2.5	1.3	1.0–1.5	1.8	1.6–2.0
Mother's education, y										
<12 vs 12–15	1.0	0.9–1.1	1.0	0.9–1.1	1.0	0.8–1.1	1.0	0.9–1.1	1.1	0.9–1.2
16+ vs 12–15	0.7	0.6–0.7	0.7	0.6–0.7	0.9	0.7–1.0	0.7	0.6–0.8	0.6	0.5–0.7
Prenatal care onset, mo										
4–6 vs 1–3	0.9	0.8–1.0	0.9	0.8–1.0	1.0	0.8–1.1	0.9	0.8–1.0	0.9	0.8–1.0
7–9 vs 1–3	0.9	0.8–1.0	0.9	0.7–1.0	1.0	0.8–1.2	1.0	0.8–1.2	0.8	0.6–1.0
Tobacco use during pregnancy, no. of cigarettes per day										
1–9	0.9	0.7–1.0	0.9	0.7–1.0	0.8	0.6–1.0	0.8	0.7–1.0	0.9	0.7–1.1
10–19	0.7	0.6–0.8	0.7	0.6–0.8	0.7	0.5–0.9	0.7	0.6–0.8	0.8	0.6–0.9
20+	0.7	0.6–0.8	0.7	0.5–0.8	0.8	0.5–1.0	0.7	0.6–0.8	0.6	0.5–0.7

Note. All factors were included in the logistic regression model simultaneously, and thus adjusted for one another.

chronic hypertension, preeclampsia, and other forms of PIH that makes comparisons among studies difficult. Several papers have recently proposed standardized classification systems,[4-6] but at present these more refined classifications are not available for large populations.

The literature with which the patterns of PIH found in North Carolina can be compared is quite limited. An examination of patterns of hospitalized cases of preeclampsia found increased risk with young maternal age (in contrast to the opposite pattern reported here), although similar risks for Blacks and Whites were found in both studies.[16] The well-recognized increased risk of preeclampsia for nulliparas and plural gestations[7] was clearly identified in our data. Chesley cites "conventional wisdom" that Blacks are at greater risk of preeclampsia than Whites and that lower social class women are at greater risk than upper social class women, which he challenges for the absence of evidence.[7] The present data indicate the absence of such racial or social class differences, with a suggestion only of lower risk among women who are neither Black nor White.

The inverse association between cigarette smoking and PIH has been reported previously.[8,9] Marcoux et al. separated preeclampsia from other forms of PIH and found the strongest inverse association with preeclampsia, with risks of 0.8, 0.6, and 0.4 for smokers of less than 11, 11 to 20, and 20 or more cigarettes per day.[9] The present study found more modest reductions in risk of 0.9, 0.8, and 0.7 for similar smoking categories but combining all forms of PIH. Generally, the results of the two studies are similar given the weaker smoking gradient found by Marcoux et al. for PIH other than preeclampsia. The quality of smoking data on birth certificates is questionable, given the stigma associated with smoking during pregnancy, but nondifferential misclassification seems most probable with a resulting bias toward the null.[17] If the reporting were biased, one would expect overreporting for women with adverse outcomes (i.e., PIH) relative to women with uncomplicated pregnancies, opposite to the observed inverse association.

The limitations in the data should be acknowledged. The principal concern is in the undocumented process by which diagnoses of PIH are made. The process by which a woman comes to be documented on the birth certificate as having PIH involves many steps, including her entry into prenatal care, diagnosis by a physician, notation in the medical record, and transfer of that information to the birth certificate. Each step is potentially fallible, with underascertainment likely. Nonetheless, the overall risk of reported PIH seems reasonable, and the corroboration of previously recognized patterns by parity and cigarette smoking provides some reassurance regarding the quality of the data.

In spite of the virtually universal adoption of the new birth certificate format, there has been little work done to empirically assess the validity of its contents. To our knowledge, the only directly applicable study was a comparison in the state of Missouri between the risk of PIH based on the birth certificate item and the risk based on hospital discharge data (W. Schramm, written communication, May 1991). There were 42.5 cases per 1000 births based on hospital discharge data vs 36.1 cases per 1000 births based on the birth certificate information, resulting in a ratio of 0.85. This would suggest modest underascertainment on birth certificates, but the actual sensitivity and specificity of classification could not be estimated. Research is badly needed to evaluate the quality of specific items on the certificate so that researchers and health care planners can make appropriate use of what will soon be widely available data. Studies of pregnancy and delivery complications other than PIH in Washington State indicate that the check-box format yields more complete reporting than an open-ended format.[18]

A second limitation of this study is that it was not possible to separate PIH into true preeclampsia (hypertension with proteinuria or edema) and other forms of PIH for separate analyses. It was, however, possible to remove women with re-

ported chronic hypertension from the analysis. To the extent that the subtypes of PIH have distinct etiologies, the results reported here are diluted.

Finally, there was a limited array of information available from the birth certificate for analysis and, for at least some of the items, questions about data quality arise. Hypotheses regarding the role of nutrition and contraception, for example, could not be addressed at all. Alcohol use was judged to be inadequately reported for analysis, based on the rarity of reported use. Many of the demographic characteristics (mother's age, education, parity, etc.) are thought to be well reported. Although tobacco smoking as reported on the birth certificate has not been validated, a study in Missouri suggested that the aggregate levels are concordant with those obtained in the National Natality Survey.[19,20] The sample included in the National Natality Survey differs markedly from the North Carolina population with regard to race and marital status, but some reassurance of reasonable data quality is provided by the observation of similar patterns in the two sources (including comparable overall prevalence, reduced smoking during pregnancy with advancing age, and a marked reduction with increasing levels of education). Furthermore, recent analysis of North Carolina birth certificate data[21] revealed the usual pattern of reduced birthweight among smokers, providing additional assurance about the quality of that item.

The pattern of results supports further epidemiologic evaluation of this highly prevalent and etiologically mysterious condition. Perhaps the associations not found are as important as those that were: race and social class were virtually unrelated to risk except for a reduced risk among women of "other" races. This might suggest the absence of environmental influences, but the clear patterns by maternal age, parity, and tobacco use indicate otherwise. Weaker associations were found with season of birth (higher risks for winter and spring births) and increased risk with the most rural residence. With the recent availability of information on PIH in vital records, broader geographic areas should be examined to develop refined etiologic hypotheses in parallel with more focused studies of existing hypotheses. □

Acknowledgments

We would like to thank Paul Buescher, PhD, for his assistance in obtaining and interpreting the data and for his thorough and constructive review of the manuscript. We would also like to acknowledge the advice on disease classification provided by Watson Bowes III, MD.

References

1. Lindmark G, Lindberg B, Hogstedt S. The incidence of hypertensive disease in pregnancy. *Acta Obstet Gynecol Scand Suppl.* 1984;118:29–32.
2. Sullivan JM. *Hypertension in Pregnancy.* Chicago, Ill: Year Book Medical Publishers; 1986:3–7.
3. Davies AM, Dunlop W. Hypertension in pregnancy. In: Barron SL, Thomson AM, eds. *Obstetrical Epidemiology.* London: Academic Press. 1983:167–208.
4. Davey DA, MacGillivray I. The classification and definition of the hypertensive disorders of pregnancy. *Am J Obstet Gynecol.* 1988;158:892–898.
5. Lindheimer MD, Katz AI. Hypertension in pregnancy. *N Engl J Med.* 1985;313:675–680.
6. Anonymous. Classification of hypertensive disorders of pregnancy. *Lancet.* 1989;1:935–936.
7. Chesley LC. History and epidemiology of preeclampsia-eclampsia. *Clin Obstet Gynecol.* 1984;27:801–820.
8. Duffus GM, MacGillivray I. The incidence of pre-eclamptic toxaemia in smokers and non-smokers. *Lancet.* 1968;1:994–995.
9. Marcoux S, Brisson J, Fabia J. The effect of cigarette smoking on the risk of preeclampsia and gestational hypertension. *Am J Epidemiol.* 1989;130:950–957.
10. Rabinowitz M, Bellinger D, Leviton A, Needleman H, Schoenbaum S. Pregnancy hypertension, blood pressure during labor, and blood lead levels. *Hypertension.* 1987;10:447–451.
11. Eskenazi B, Bracken MB, Holford TR, Grady J. Exposure to organic solvents and hypertensive disorders of pregnancy. *Am J Ind Med.* 1988;14:177–188.
12. Marcoux S, Brisson J, Fabia J. The effect of leisure time physical activity on the risk of pre-eclampsia and gestational hypertension. *J Epidemiol Community Health.* 1989;43:147–152.
13. Campbell DM, MacGillivray I. Pre-eclampsia in second pregnancy. *Br J Obstet Gynecol.* 1985;92:131–140.
14. Belizan JM, Villar J, Repke J. The relationship between calcium intake and pregnancy-induced hypertension: up-to-date evidence. *Am J Obstet Gynecol.* 1988;158:898–902.
15. Klonoff-Cohen H, Savitz DA, Cefalo RC, McCann MF. An epidemiologic study of preeclampsia. *JAMA.* 1989;262:2143–2147.
16. Saftlas AF, Olson DR, Franks AL, Atrash HK, Pokras R. Epidemiology of preeclampsia and eclampsia in the United States, 1979–1986. *Am J Obstet Gynecol.* 1990;163:460–465.
17. Rothman KJ. *Modern Epidemiology.* Boston, Mass: Little, Brown; 1986.
18. Frost F, Starzyk P, George S, McLaughlin JF. Birth complication reporting: the effect of birth certificate design. *Am J Public Health.* 1984;74:505–506.
19. Stockbauer JW, Land GH. Changes in characteristics of women who smoke during pregnancy: Missouri, 1978–88. *Public Health Rep.* 1991;106:52–58.
20. Prager K, Malin H, Spiegler D, Van Natta P, Placek PJ. Smoking and drinking behavior before and during pregnancy of married mothers of live-born infants and stillborn infants. *Public Health Rep.* 1984;99:117–127.
21. Buescher PA. *New Items on the North Carolina Birth Certificate.* Raleigh, NC: Division of Statistics and Information Services, Center for Health and Environmental Statistics, North Carolina Department of Environment, Health, and Natural Resources; 1991. CHES Studies Report No. 58.

The Transformation of American Midwifery: 1975 to 1988

Eugene R. Declercq, PhD

ABSTRACT

Background. The use of midwives is a natural solution to the problem of improving access to skilled perinatal services while lowering costs. The number of midwife-attended births has grown from 0.9% of all births in 1975 to 3.4% of all births in 1988. The purpose of the study was to determine how mothers served by midwives and the settings in which they are served have changed in that period.

Methods. The analysis is based on birth certificate data from 1975 to 1988 from the Natality, Marriage and Divorce Statistics Branch of the National Center for Health Statistics, Centers for Disease Control.

Results. Almost all of the growth (93.2%) in midwife-attended births from 1975 to 1988 was in hospitals; 87.3% of all births attended by midwives occurred in hospitals. Pronounced differences exist between mothers served by midwives in and outside of hospitals, and there are strong regional patterns in midwife attendance at birth.

Conclusions. Given the positive outcomes associated with midwifery practice, further research into the content of midwifery care is recommended. (*Am J Public Health.* 1992; 82:680–684)

Introduction

The practice of midwifery in the United States underwent certain changes between 1975 and 1988. (1975 is the first year for which data permit a detailed examination of midwife-attended births.) Midwives seem a natural solution to the problems of escalating costs and limited access in the US health care system.[1–3] Despite a heavy reliance on midwifery in many European countries with positive perinatal outcomes,[4–6] the use of midwives in the United States has been limited for a variety of reasons,[7] although there has been a significant recent growth in the number of nurse-midwives.[8] The last major national study to examine comparable data[9] was based on findings from 1978 and 1979, and several notable changes have been documented since that study. In this paper I examine the characteristics of mothers served by midwives, the settings in which midwives attend births (in or outside of hospitals), and regional variations in these factors.

Methods

All results presented are based on birth certificate data provided by the Natality, Marriage and Divorce Statistics Branch of the National Center for Health Statistics, Centers for Disease Control. With the exceptions of mother's education (73%) and Apgar scores (77%), the variables studied were reported in at least 99% of the cases. The basic comparisons will distinguish midwife-attended births, both in and outside of hospitals, from all births in the United States.

Results

Contemporary Midwifery Practice

Figure 1 illustrates the substantial growth in midwife attendance at births over the years from 1975 to 1988. Virtually all (93.2%) of the increase occurred in hospital births.[10] Overall, the proportion of all births that were attended by midwives increased from 0.9% in 1975 to 3.4% in 1988[10]; the number of midwife-attended births outside of hospitals decreased from 33.1% of all midwife-assisted births in 1975 to 12.7% in 1988. Because the birth certificate category of "midwife" as attendant includes both nurse-midwives and lay midwives, and the latter virtually never have hospital privileges,[11] the concentration of midwife practice in hospitals points to the growing dominance of nurse-midwives. These distinctions should become clearer with 1989 birth certificate data, which will, for the first time, distinguish between lay midwives and nurse-midwives.

Past research has found distinctions between the populations served by midwives in and outside of hospitals[9]; therefore, in the results shown in Table 1, both attendant and place of delivery have been controlled for. Midwives in hospitals serve mothers who are distinctly younger than average (48.3% are younger than 25 years, compared with 39.8% of all mothers), whereas midwives outside of hospitals at-

The author is with the Merrimack College Political Science Department, North Andover, Mass.

Requests for reprints should be sent to Eugene R. Declercq, PhD, Merrimack College, Department of Political Science, North Andover, MA 01845.

This paper was submitted to the Journal April 2, 1991, and accepted with revisions November 27, 1991.

tend mothers who are older than average (40.2% are older than 30 years, compared with 28.5% of all mothers). Midwife-attended births in hospitals differ somewhat from the norm in terms of proportion nonwhite. Mothers attended by midwives out of hospital are much more likely (94.6%) to be White. Midwives as a whole are also much more likely to serve foreign-born mothers (13.5% of all births are to foreign born mothers compared to 22.8% of midwife attended births); a significant proportion of those attended by midwives are Mexican-born mothers in Texas. Texas alone accounts for one third of all midwife-attended out-of-hospital births in the United States.

The findings on parity mirror those concerning age. Midwives attend mothers of much higher parity out of hospital (47.4% are parity 3+) than in hospital (29.4% are parity 3+) or all births (26.5%). Four states (California, Texas, Washington, and New York [outside of New York City]) do not report data on the education of mothers; therefore, the analysis of maternal education must be limited. However, the trends noted elsewhere are reaffirmed: mothers who are attended by midwives in hospitals are less well educated than average, and mothers who are attended by midwives outside of hospitals are generally better educated. The distinctions between hospital and out-of-hospital practice are even clearer with regard to marital status: 35.4% of midwife-attended hospital births are to unmarried mothers, and only 9.7% of out-of-hospital births attended by midwives are to unmarried mothers. This is a partial reflection of the difference in proportion of black mothers served in each setting.

Midwife-attended mothers typically begin prenatal care later than average, particularly in the case of out-of-hospital births. What is unknown, and clearly needs further study, is both the provider and the content of prenatal care delivered in these various settings. Mothers attended by midwives, both in and outside of hospitals, have low-birthweight babies (< 2500 g) at less than half the general rate. This is particularly true for babies born outside of hospitals, which is not surprising, given the characteristics of mothers who give birth outside of hospitals and past studies on birthweight and Apgar scores in out-of-hospital births.[12] There is little variance in low Apgar scores (< 7) among any of the comparison groups. Of the babies born out of hospital to mothers attended by midwives, 58.5% have a 1-minute Apgar score of 9 or 10. Of all births, 42.1% of babies achieve such a score; for midwife-attended hospital births, the figure is 45%. These findings are at least in part a function of careful risk assessment by midwives and referral of higher risk cases to doctors. The extent to which the nature of midwifery care plays a role in these results remains unclear.

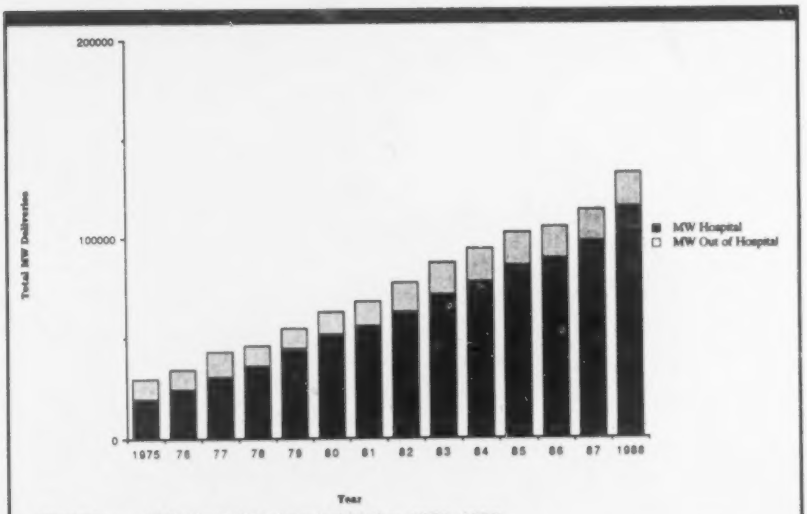


FIGURE 1—Midwife (MW)-attended births, 1975–1988.

TABLE 1—Percentage Distribution of All Births and Midwife-Attended Births by Selected Characteristics, 1988

	All Births	Midwife in Hospital	Midwife Out of Hospital
Mother's age			
<19	12.5	17.8	6.3
20–24	27.3	30.4	22.3
25–29	31.7	27.5	31.4
30–34	20.6	17.3	26.7
35–39	6.9	6.0	11.6
40+	1.0	0.8	1.9
Parity			
1	41.0	39.0	25.5
2	32.7	31.9	30.5
3	16.2	16.9	22.7
4+	10.0	12.5	24.7
Mother's education, y			
0–11	20.5	30.9	21.4
12	41.1	37.7	26.3
13–15	20.8	17.1	24.7
16+	17.6	14.3	25.6
Month prenatal care began			
1–2	54.9	39.3	29.7
3	21.1	23.4	25.9
4–6	18.0	27.3	31.5
7–9 & no care	6.1	10.1	12.9
Birthweight, g			
<2500	6.9	3.4	2.0
2500–2999	18.0	15.8	10.0
3000–3499	36.6	39.6	32.5
3500–3999	29.4	30.8	35.8
4000+	11.1	10.4	19.9
1-Minute Apgar score			
<7	8.9	7.4	7.3
7–8	49.0	47.6	34.2
9–10	42.1	45.0	58.5

State Analysis

Because the laws regulating midwives vary widely across states,[11,13] I an-

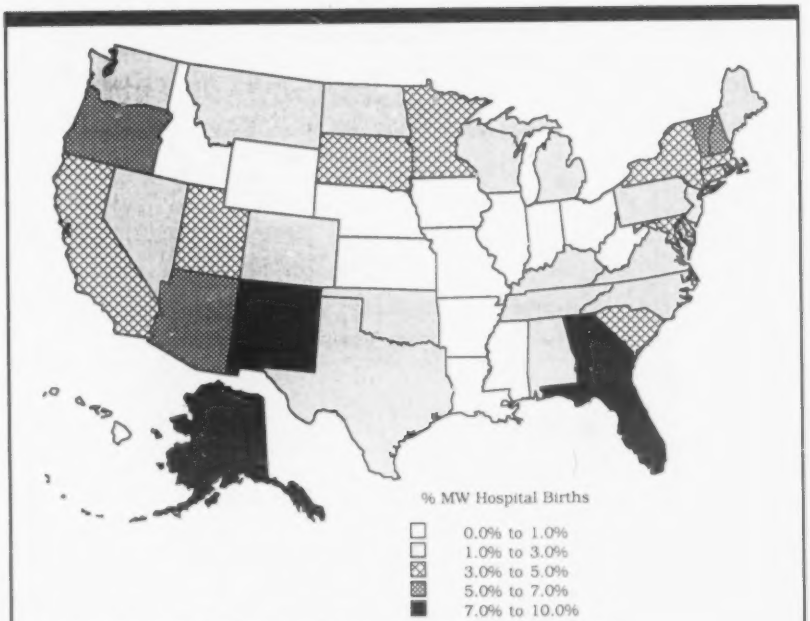


FIGURE 2—Midwife (MW)-attended hospital births by state, 1988.

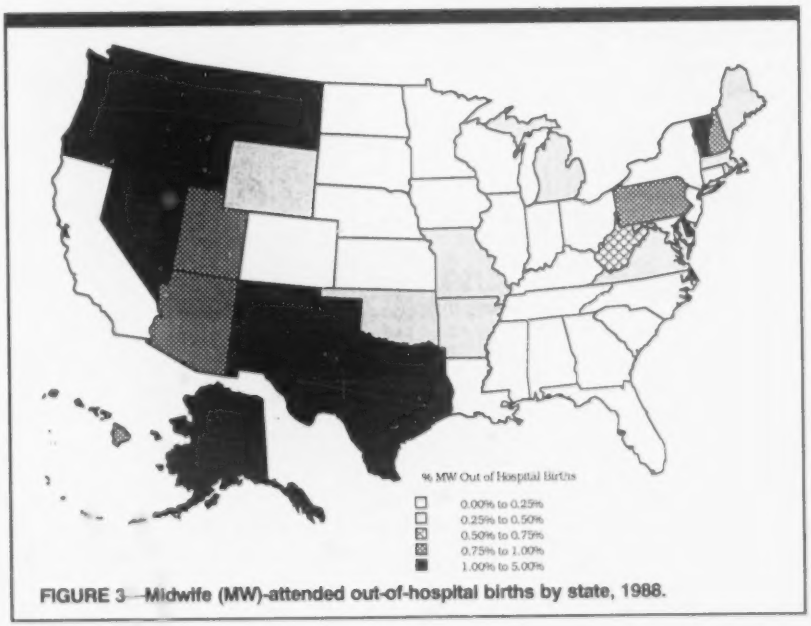


FIGURE 3—Midwife (MW)-attended out-of-hospital births by state, 1988.

alyzed interstate differences. Figure 2 illustrates regional variations in midwife-attended hospital births in 1988. Midwife attendance at hospital births was concentrated along the coasts and in the upper Midwest. Figure 3 also shows a strong regional pattern in the prevalence of midwife-attended out-of-hospital births: it is largely a Western phenomenon. Eight of the nine states where midwife-attended out-of-hospital births account for more than 1% of all births are located west of the Mississippi. In only four Eastern states are more than 0.5% of all births out-of-hospital births attended by a midwife. In this case there appears to be a potential inverse relationship with urbanization.

The greatest change in midwife-attended births over the years from 1975 to 1988 was concentrated, not surprisingly, in those states where hospital-based midwife activity was dominant (Figure 4). The 15 states that experienced at least a 4-percentage-point increase are located at the four corners of the continental United States; some (Florida, Connecticut, Minnesota, Alaska, and California) clearly show a change based on the increase in midwife-attended hospital births, whereas others (Vermont, Oregon) have a more mixed pattern. Four of the five states that experienced a decrease were in the South, and the decline was concentrated among Black mothers.

The analysis suggested a number of variables that might influence the frequency of midwife-attended deliveries in a state (e.g., population density, wealth, race) and there are several variables of interest (e.g., neonatal and infant mortality rates) that might in turn be related to midwife activity. An analysis of the Pearson product-moment correlations between a range of relevant variables and midwife-attended births in the states uncovered few statistically significant relationships. Of interest is the lack of a stronger relationship between midwife activity in 1975 and in 1988, with a correlation of only .32 between the proportions of all births in a state attended by midwives. Therefore, not only was there growth in the number of midwife-attended births, but there was a changing pattern in terms of the states where midwifery was popular.

Changes in Midwifery Practice, 1975–1988

Perhaps the most noteworthy change in midwifery practice in the period studied was in the racial composition of the mothers served by midwives in 1988. Although births to non-White mothers as a proportion of all births increased from 1975 to 1988, the proportion of non-White mothers served by midwives decreased dramatically, particularly in the out-of-hospital population (Figure 5). The pattern among those non-white mothers who are Black is the same, but with an even more precipitous decline.

Differences can also be seen between Taffel's[9] analysis of data for 1978 and 1979 and the 1988 data on some of the demographic variables discussed above. Maternal characteristics, such as parity, marital status, and prenatal care, show little change over the decade. Age differences between the populations served by midwives and mothers in general were more pronounced in 1988 than in 1978–1979. Although the proportion of mothers giving birth after the age of 30 rose in general over the decade (from 19% to 29%), the growth was even greater among midwife-attended mothers giving birth outside of

hospitals (from 23% to 40%). The reverse occurred in the case of nativity. An increase in the proportion of foreign-born mothers in general (from 10% to 14%) and a decrease in the proportion of midwife-attended mothers delivering out of hospital who were foreign-born (from 29% to 25%) resulted in differences in nativity being considerably lessened. Although the overall incidence of low birthweight did not change over the decade, the proportion of low-birthweight infants born to mothers who were attended by midwives in hospitals decreased from 4.9% in 1978–1979 to 3.4% in 1988; for mothers attended by midwives outside of hospitals, the proportion decreased from 3.8% in 1978–1979 to 2.0% in 1988. This decrease in low-birthweight babies born to mothers attended by midwives is partly the result of the decrease in the number of Black mothers served by midwives. Overall, the changes in midwifery practice between 1978–1979 and 1988 appear to have further distinguished out-of-hospital births from those in hospitals: out-of-hospital midwife-attended births seem concentrated among older White mothers.

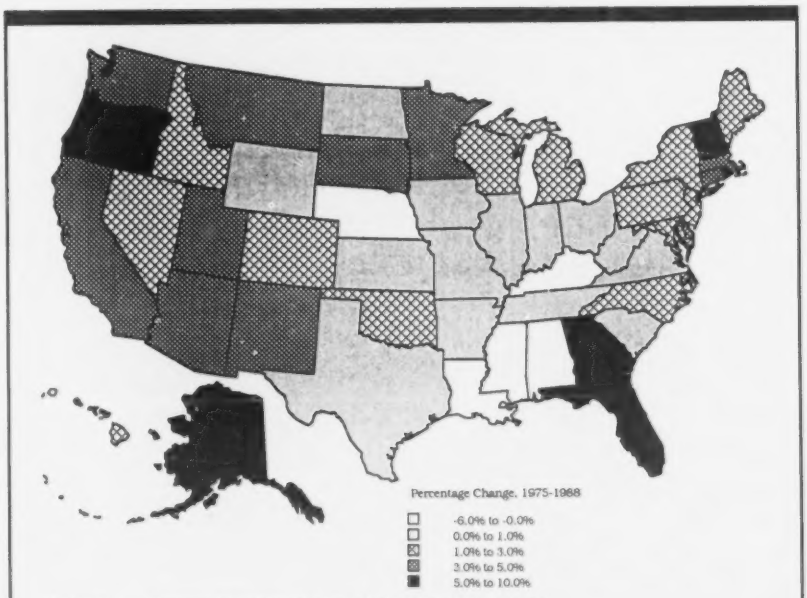


FIGURE 4—Change in midwife-attended births, 1975–1988.

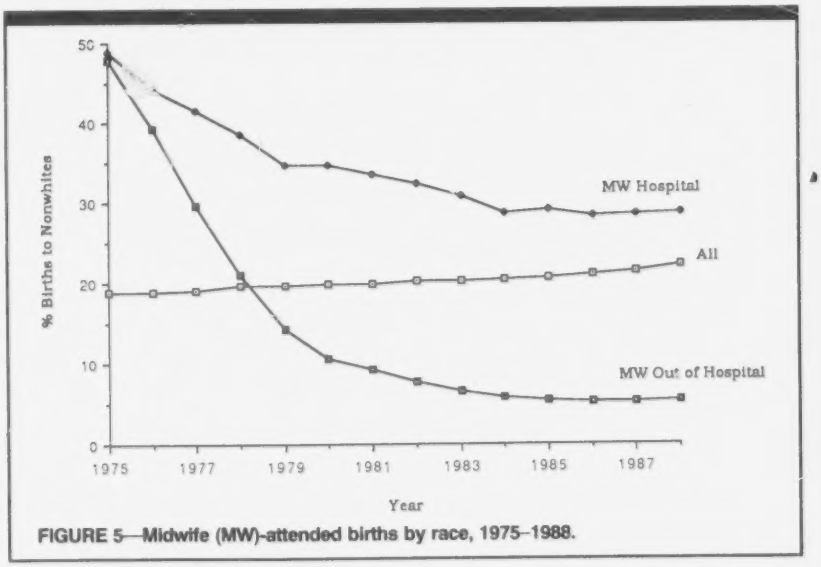


FIGURE 5—Midwife (MW)-attended births by race, 1975–1988.

Discussion

How can contemporary midwifery practice be summarized? Midwifery practice in the United States is currently hospital based: 87.3% of all midwife-attended births occurred in a hospital in 1988. Midwives who practice in hospitals serve mothers who are younger; who are more likely to be unmarried, more likely to be foreign-born, and more likely to be minorities; and who receive less prenatal care than the average mother in the United States. Midwives who attend mothers outside of hospitals support a clientele that is older, has higher parity, and is much more likely to be White; is more likely to be foreign-born and married; and receives less prenatal care than mothers as a whole. Judged by the only two measures of outcome available on birth certificates, birthweight and Apgar score, mothers and babies have distinctly better than average outcomes when births are attended by midwives, either in or out of hospitals. Despite the obvious limitations of birthweight and Apgar scores as measures of outcome, when one considers that midwives, particularly in hospitals, are serving a population apparently at risk, something about the system of care developed in these cases (careful screening of risk with referral of high risk cases to doctors, more attention to prenatal care, less intervention at birth?) is working.

The findings of positive outcomes to midwifery care are hardly unique to this study, and recent research on birth centers[14] provides more support for the increased use of midwives and suggests the need for systematic studies of the cost savings associated with midwifery care. The degree to which some of the different patterns noted above are a result of funding mechanisms (e.g., Medicaid reimbursement rules) also deserves further inquiry. Midwives can help relieve problems of access for poor and minority mothers, and midwifery care in hospitals is increasingly serving these populations. Pressures to concentrate midwifery care in these settings will likely grow and result in a clear split in midwifery practices between midwives serving the poorest mothers in hospitals and those serving upper-middle-class mothers outside of hospitals.

I found a strong regional pattern of midwife attendance at birth, and distinct interstate differences both at present and in developments over time. Because midwifery is regulated at the state level, differences are likely to persist, permitting the systematic study of the impact of both ex-

isting policies and innovative approaches to the further integration of midwifery into perinatal care in the United States. □

Acknowledgments

An earlier version of this paper was presented at the 118th Annual Meeting of the American Public Health Association, New York, NY, October 1990. This research was supported by a grant from the Merrimack College Faculty Development Committee.

References

1. Levit K, Freeland M, Waldo D. National health care spending trends: 1988. *Health Affairs.* 1990;9:171–184.
2. Nesbitt T, Connell F, Hart L, Rosenblatt R. Access to obstetric care in rural areas: effect on birth outcomes. *Am J Public Health.* 1990;80:814–818.
3. Rosenbiatt R. The perinatal paradox: doing more and accomplishing less. *Health Affairs.* 1989;8:158–168.
4. Kitzinger S. *The Midwife Challenge.* London: Pandora; 1988.
5. Miller CA. *Maternal Health and Infant Survival.* Washington, DC: National Center for Clinical Infant Programs; 1987.
6. Wagner M. Infant mortality in Europe: implications for the United States. *J Public Health Policy.* 1988;9:473–484.
7. Litoff J. *American Midwives: 1860 to the Present.* Westport, Conn: Greenwood Press; 1978.
8. Lehrman E, Paine L. Trends in nurse-midwifery. *J Nurse Midwifery.* 1990;35:192–203.
9. Taffel S. *Midwife and Out-of-Hospital Deliveries, United States.* Washington, DC: US Government Printing Office; 1984. *US Dept of Health and Human Services* publication PHS 84–1918.
10. National Center for Health Statistics. *Advance Report of Final Natality Statistics, 1988.* Hyattsville, Md: US Public Health Service; 1990. Monthly Vital Statistics Report 39(Suppl).
11. Butter I, Kay B. State laws and the practice of midwifery. *Am J Public Health.* 1988; 78:1161–1169.
12. Declercq E. Birthweight and Apgar scores as measures of outcomes in out of hospital births in the U.S. *Public Health Rep.* 1984;99:63–73.
13. Mullinax K. Supplemental report on nurse-midwifery legislation. *J Nurse Midwifery.* 1987;32:156–159.
14. Rooks J, Weatherby N, Ernst E, Stapleton S, Rosen D, Rosenfield A. Outcomes of care in birth centers. *N Engl J Med.* 1989; 321:1804–1811.

Errata

In: Rigotti NA, Bourne D, Rosen A, Locke JA, Schelling TC. Workplace compliance with a no-smoking law: a randomized community intervention trial. *Am J Public Health.* 1992;82:229-235.

This study should be attributed to the Institute for the Study of Smoking Behavior and Policy, John F. Kennedy School of Government, Harvard University.

In: Cohen D, Scribner R, Clark J, Cory D. The potential role of custody facilities in controlling sexually transmitted diseases, *Am J Public Health.* 1992;82:552-556.

The authors' professional associations at the time of the study were as follows: For nearly the entire project, Deborah Cohen was with the University of Southern California, acting as a consultant for the Los Angeles County Department of Health Services; Richard Scribner and David Cory were with the Los Angeles County Department of Health Services. The authors' current associations are as follows: Deborah Cohen and Richard Scribner are with the University of Southern California, and David Cory is an independent consultant. (John Clark's affiliation with the Los Angeles County Sheriff's Department is unchanged.)

The following disclaimer should have been included: "The views expressed in the article are those of the authors and may not represent the views of the Los Angeles County Department of Health Services."

In: Jemmott JB III, Jemmott LS, Fong GT. Reduction in HIV risk-associated sexual behaviors among black male adolescents: effects of an AIDS prevention intervention. *Am J Public Health.* 1992;82:372-377.

In Table 2, the 95% Confidence Intervals of Difference for Attitudes and Intentions were misaligned. The corrected table is printed here:

TABLE 2—Adjusted[a] 3-Month Follow-up Means by Experimental Condition

Variable	AIDS Prevention Condition	Control Condition	Difference	95% CI of Difference
Knowledge	47.20 (83)	44.40 (67)	.28	.72, 4.88
Attitudes	3.13 (82)	3.38 (67)	−.25	−.54, .04
Intentions	2.87 (82)	3.30 (67)	−.43	−.74, −.12
Risky behavior	−.12 (76)	.24 (62)	−.36	−.64, −.08

Note. Numbers of respondents are shown in parentheses. CI = confidence interval. Higher numbers indicate greater AIDS knowledge, more positive attitudes toward behaviors that increase risk of sexually transmitted HIV infection, stronger intentions to engage in such risky behaviors, and more reports of risky sexual behavior.
[a]For each variable, the preintervention measure is partialled out of the 3-month follow-up measure.

The Influence of Social and Political Violence on the Risk of Pregnancy Complications

B. Cecilia Zapata, DrPH, Annabella Rebolledo, PhD, Eduardo Atalah, MD, Beth Newman, PhD, and Mary-Claire King, PhD

ABSTRACT

Background. Events in Chile provided an opportunity to evaluate health effects associated with exposure to high levels of social and political violence.

Methods. Neighborhoods in Santiago, Chile, were mapped for occurrences of sociopolitical violence during 1985–86, such as bomb threats, military presence, undercover surveillance, and political demonstrations. Six health centers providing prenatal care were then chosen at random: three from "high-violence" and three from "low-violence" neigborhoods. The 161 healthy, pregnant women due to deliver between August 1 and September 7, 1986, who attended these health centers were interviewed twice about their living conditions. Pregnancy complications and labor/delivery information were subsequently obtained from clinic and hospital records.

Results. Women living in the high-violence neighborhoods were significantly more likely to experience pregnancy complications than women living in lower violence neighborhoods (OR=5.0; 95% CI=1.9–12.6; p<0.01). Residence in a high-violence neighborhood was the strongest risk factor observed; results persisted after controlling for several sets of potential confounders.

Conclusion. Living in areas of high social and political violence increased the risk of pregnancy complications among otherwise healthy women. (*Am J Public Health.* 1992; 82:685–690)

Introduction

Research on the health effects of social and political violence has typically focused on individuals victimized by violent acts[1–8] and neglected the population at large, which in a troubled time or place is often exposed to a pervasive atmosphere of fear and tension. The influence of such an atmosphere on health is now being investigated more frequently.[9,10] The recent political unrest in Chile provided an opportunity to assess the impact of social and political violence on the risk of pregnancy complications in otherwise healthy women.

Between 1973, when President Salvador Allende was overthrown by a military coup d'état, and 1990, Chile experienced widespread social and political upheaval. The commanders-in-chief of the Army, Navy, Air Force, and National Police created a military junta, headed by General Augusto Pinochet Ugarte. The junta adjourned the National Congress, dismissed elected officials, closed the constitutional court, and banned activities of political parties and labor unions. Extrajudicial killings, disappearances, forcible exiles, and arbitrary arrests reached unprecedented levels.[11,12] The government rationalized its activities by decreeing a "state of exception" in 1973, which permitted new legislation restricting basic human rights.[11] Under this decree, the military junta declared the country variously under state of siege, state of emergency, and state of alert.

During the period of this study, from December 1985 to December 1986, Chile was continually under a state of exception or state of emergency (December 1985 to August 1986) or under a state of siege (September to December 1986). The atmosphere in Chile, especially in greater Santiago, was one of confrontation between civilians and the police and armed forces. During this period, there were 33 665 political arrests.[13,14] An especially active period was April 29 to May 13, 1986, when joint military and police forces raided 33 neighborhoods in greater Santiago and southern Chile. Despite the government's prohibition of strikes, Chileans went on a nationwide general strike on July 2 and 3, 1986. The government declared the strike illegal and hundreds of people were arrested. In Santiago 2 people were burned alive by the military (one died, the other survived),[14] 4 opponents of the government were kidnapped and executed, and at least 132 cases of torture were openly denounced in the courts. The staff of the Human Rights Commission suffered 32 arrests, 2 kidnappings, 86 threats, and 13 cases of cruel, inhumane, or degrading treatment.[13] The Vicariate of Solidarity of the Catholic Church had 5 of its staff arrested, received 17 threats, and experienced 1 unsuccessful assassination attempt.[13] A physician and a lawyer employed by the Vicariate, as well as a private physician, faced charges of terrorism

At the time of this study, B. Cecilia Zapata, Beth Newman, and Mary-Claire King were with the School of Public Health, University of California, Berkeley. B. Cecilia Zapata and Beth Newman are now with the School of Public Health, University of North Carolina at Chapel Hill. Annabella Rebolledo and Eduardo Atalah are with the Faculdad de Medicina Norte, Universidad de Chile, Santiago.

Requests for reprints should be sent to B. Cecilia Zapata, DrPH, Department of Maternal and Child Health, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7400.

This paper was submitted to the Journal May 28, 1991, and accepted with revisions December 31, 1991.

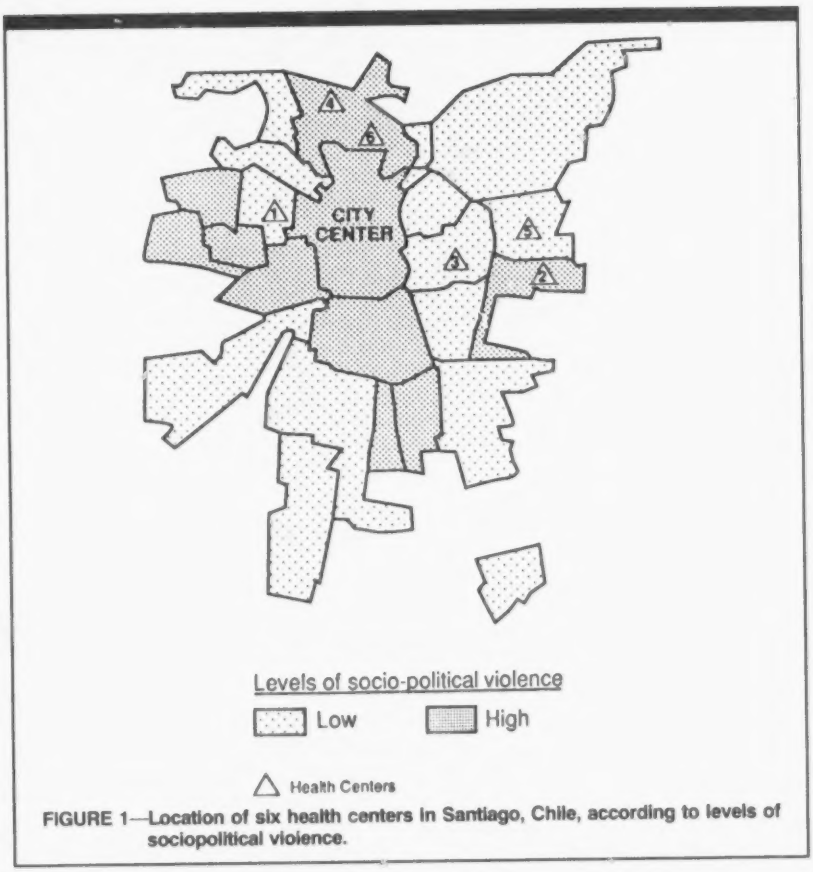


FIGURE 1—Location of six health centers in Santiago, Chile, according to levels of sociopolitical violence.

because they assisted a wounded man who sought medical assistance in a clinic under contract with the Vicariate of Solidarity.[13]

However, social and political violence, unrest, and crime were not distributed equally throughout Chile in general, or across Santiago in particular. Our goal was to study prospectively pregnancy complications among healthy Chilean women who received prenatal care at the health centers but who lived under different levels of social and political violence.

Methods

For the purpose of this study, a team of researchers from the School of Medicine of the Universidad de Chile mapped Santiago according to the frequency of occurrence of five classes of events: police and military presence, undercover police activity, bomb threats, demonstrations, and lack of reliable public transportation before the curfew hour. Neighborhoods were classified as having high or low levels of sociopolitical violence (Figure 1). From among Santiago's 58 health centers, we randomly selected 3 from high-violence areas and 3 from low-violence areas. The director of each of these six health centers was then asked to independently rate the level of sociopolitical violence in the neighborhood of his health center. The directors' assessments were generally in agreement with our mapping designations. However, some of the directors pointed out small areas that were within generally low-violence neighborhoods, but because of special circumstances were actually at high risk of sociopolitical violence (e.g., a block containing the heavily guarded home of a member of the military junta). The 22 women who lived in such pockets were reclassified from the low-violence to the high-violence category to reflect actual exposure. Assignments to exposure were made prior to any data collection.

Our study sample comprised pregnant women between 19 and 40 years of age who were enrolled for prenatal care at the six selected health centers, lived within the clinics' target areas, had no chronic medical conditions, and (based on the date of the last menstrual period) were due to deliver between August 1 and September 7, 1986. The sample consisted of 179 women, 113 living in areas with high levels of sociopolitical violence and 66 living in areas of lower sociopolitical violence.

The first interview with each woman was conducted between March 13 and April 13, 1986, by professional interviewers from the Universidad de Chile and one of the authors (B.C.Z.). The interviews took place in the subjects' homes. The second interviews took place from June 18 through July 7, 1986. These dates correspond to the end of the first trimester or the beginning of the second trimester of pregnancy for the first interview and to the third trimester for the second interview. Between the times of the first and second interviews, 13 respondents moved out of the target area, 4 had late miscarriages, and 1 refused to continue to participate for lack of time. The final sample was therefore composed of 161 women, 106 from areas of high sociopolitical violence and 55 from areas of low sociopolitical violence. The 18 women lost to the second interview did not differ from the final cohort with respect to age, marital status, or prenatal care.

The interviews addressed potential influences on pregnancy, including length of residence in the neighborhood, the respondent's positive and negative relationships with her neighbors, family composition and relationships between the respondent and members of the household, strains within the family, violence in the home, employment and personal economic status, smoking and alcohol consumption, nutrition, health-related issues, and educational level. A series of questions allowed the construction of scales measuring social support,[15] life events,[16] alienation,[17] depression,[15,16,18,19] locus of control (helplessness),[15–17] and neighborhood satisfaction. All scales were coded so that a higher score represented a potentially negative attribute, for example, more negative life events or less social support. We also asked the women a series of questions about their perceptions of sociopolitical violence in their neighborhoods (neighborhood milieu) at interviews 1 and 2.[20]

All women received prenatal and delivery care from professional midwives. If a woman developed a health problem that might influence her pregnancy or delivery, she was referred to an obstetrician. All diagnoses of pregnancy and delivery complications were made by this team of health care providers after appropriate clinical and laboratory tests and were documented in the medical records. Medical

records were reviewed at the health centers after the second interview was completed, and hospital records were reviewed after delivery. All medical records were abstracted according to a structured abstracting form by three health professionals from the School of Medicine of the Universidad de Chile who did not know exposure status. Complications of pregnancy were defined as any of the following clinical diagnoses: gestational hypertension; pre-eclampsia, eclampsia, or toxemia; bleeding at first, second, or third trimester; threat of miscarriage; threat of premature labor; premature rupture of membranes; late miscarriage (stillbirth); painful preterm contractions requiring bed rest and sexual abstinence or hospitalization; and fetal growth retardation. A woman was considered subject to pregnancy complications if she had any two or more of these complications or one extremely severe complication. In the latter category were three women, two who developed eclampsia and one who gave birth to a stillborn fetus.

Odds ratios (ORs) and confidence intervals (CIs), as well as chi-square and *t* test statistics, were used to test associations of possible risk factors and pregnancy complications. Risk factors associated with pregnancy complications were also entered into logistic regression models to adjust for confounding effects. SAS data management and analysis programs were used throughout.[21–23]

Results

All 161 women had prenatal care throughout their pregnancies. The women who fell below expected weight gain during pregnancy (19%) were assigned by the health care provider to a supplementary food program through the health center. Pregnancy complications by health center are indicated in Table 1. Forty-six women had two or more complications; 115 women had one or zero complications. These two groups did not differ (at $P < .20$) in most demographic and health-related characteristics (Table 2). Age was associated with increased risk of pregnancy complications ($\chi^2 = 6.49$, $df = 3$, $P = .09$): women younger than 25 years of age and women older than 34 years of age were at greater risk of two or more pregnancy complications.

Among psychosocial indicators, risk of pregnancy complications was marginally associated with higher levels of alienation ($P = .13$) and lower levels of social support ($P = .18$) (Table 3). Other psy-

TABLE 1—Number of Women with Pregnancy Complications by Health Center, Santiago, Chile, 1986

Complication	Health Center[a] 1	2	3	4	5	6
Gestational hypertension	1	6	2	3	3	1
Toxemia/pre-eclampsia/eclampsia	1	2	2	5	1	0
Hemorrhage 1st, 2nd, 3rd trimester	0	4	3	3	4	4
Threat of miscarriage	0	4	7	6	1	3
Threat of premature labor	0	3	5	3	0	2
Premature rupture of membranes	4	7	3	2	2	5
Preterm contractions	1	14	6	14	1	6
Fetal growth retardation	0	2	6	1	0	1
Late miscarriage/stillbirth	0	1	2	1	0	0
0–1 Pregnancy complications	20	17	18	24	22	14
≥ 2 Pregnancy complications	1	12	11	12	3	7
Total pregnancies	21	29	29	36	25	21

[a]Health centers numbered 1, 3, and 5 were located in low-violence neighborhoods; those numbered 2, 4, and 6 were located in high-violence neighborhoods. Numbers in subsequent tables may differ because 22 women were reclassified from low- to high-violence exposure status (see Methods section).

TABLE 2—Distribution of Correlates of Pregnancy Complications among Women with Two or More Complications (n = 46) Compared with Women with One or No Complications (n = 115), Santiago, Chile, 1986

Correlates of Complications	Complications ≥ 2 No. (%)	0–1 No. (%)	*P* Value
Maternal age			.09
19–24 years	25 (54.3)	50 (43.5)	
25–29 years	9 (19.6)	38 (33.0)	
30–34 years	7 (15.2)	23 (20.0)	
35–40 years	5 (10.9)	4 (3.5)	
Parity			.26
0	17 (36.9)	42 (36.5)	
1–2	20 (43.5)	61 (53.0)	
≥ 3	9 (19.6)	12 (10.4)	
Gravidity			.29
Primigravida	15 (32.6)	38 (33.0)	
Secundigravida	13 (28.3)	29 (25.2)	
Multigravida	18 (39.1)	48 (41.7)	
Previous spontaneous abortion	11 (23.9)	23 (20.0)	.58
Nutritional risk factors			
Underweight	10 (21.7)	20 (17.4)	.52
Any alcohol intake	14 (30.4)	27 (23.5)	.36
Any cigarette smoking	14 (30.4)	27 (23.5)	.36
Any coffee intake	13 (28.3)	33 (28.8)	.96
Any tea intake	16 (34.8)	30 (26.1)	.27
Monthly income			.33
< US$65	28 (60.9)	59 (51.3)	
US$65–100	14 (30.4)	36 (31.3)	
> US$100	4 (8.7)	20 (17.4)	
Education			.99
1–3 years	22 (47.8)	53 (46.1)	
4–8 years	9 (19.6)	23 (20.0)	
9–10 years	10 (21.7)	27 (23.5)	
≥ 11 years	5 (10.9)	12 (10.4)	
Marital status			.89
Married/living with partner	38 (82.6)	96 (83.5)	
Single	8 (17.4)	19 (16.5)	
Work during pregnancy			.45
Working	13 (28.3)	26 (22.6)	
Not working	33 (71.7)	89 (77.4)	

TABLE 3—Distribution of Psychological Correlates of Pregnancy Complications among Women with Two or More Complications (n = 46) Compared with Women with One or No Complications (n = 115), Santiago, Chile, 1986

Correlates of Complications	Complications (mean ± SD) ≥ 2	Complications (mean ± SD) 0–1	*P* Value
Resources			
Social support (range: 12–24)	16.3 ± 3.5	15.4 ± 3.8	.18
Locus of control (range: 2–11)	5.6 ± 1.5	5.9 ± 1.7	.37
Neighborhood satisfaction (range: 3–14)	9.2 ± 2.6	8.7 ± 2.3	.20
Constraints			
Alienation (range: 0–7)	3.6 ± 1.7	3.1 ± 1.6	.13
Uncertainty (range: 0–7)	4.3 ± 1.5	4.1 ± 1.4	.45
Life events (interview 1) (range: 0–34)	10.6 ± 6.2	10.2 ± 7.1	.73
Life events (interview 2) (range: 0–25)	6.6 ± 4.6	4.7 ± 4.2	.25
Depression (interview 1) (range: 0–12)	5.7 ± 2.8	5.2 ± 2.7	.24
Depression (interview 2) (range: 0–13)	5.7 ± 2.8	5.2 ± 2.9	.36
Marital strain (range: 0–18)	4.8 ± 3.9	4.9 ± 3.6	.96

TABLE 4—Distribution of Sociopolitical Violence among Women with Two or More Pregnancy Complications (n = 46) Compared with Women with One or No Complications (n = 115), Santiago, Chile, 1986

Violence Level	Complications ≥ 2 No. (%)	Complications 0–1 No. (%)	*P* Value
In the neighborhood (classified by researchers)			<.01
High	40 (37.7)	66 (62.3)	
Low	6 (10.9)	49 (89.1)	
In the neighborhood (subject's perception at interview 1)			.17
High	38 (31.4)	83 (68.6)	
Low	8 (20.0)	32 (80.0)	
In the neighborhood (subject's perception at interview 2)			.09
High	29 (34.1)	56 (65.9)	
Low	17 (2.4)	59 (77.6)	
Violence in the home			.23
High	30 (32.3)	63 (67.7)	
Low	16 (23.5)	52 (76.5)	

TABLE 5—Independent Predictors of Pregnancy Complications among Women in Santiago, Chile, 1986: Final Multiple Logistic Regression Model

Risk Factor	Odds Ratio	95% Confidence Interval	*P* Value
Sociopolitical violence (high vs low, defined by mapping)	5.0	1.83, 12.77	<.01
Social support[a]	1.9	0.92, 4.07	.08
Neighborhood milieu at interview 1	1.9	0.78, 4.86	.15

[a]Scales were dichotomized, and odds ratios compare women scoring above the median to women scoring below the median (reference group).

chosocial indicators were not associated with risk of complications.

The fundamental question addressed in this study was whether living in a neighborhood of high sociopolitical violence was associated with pregnancy complications. Of the 55 women living in areas of low sociopolitical violence, 6 (11%) had at least two pregnancy complications, whereas among the 106 women living in areas of high sociopolitical violence, 40 (38%) had at least two pregnancy complications (Table 4). Thus women living in high-violence neighborhoods, as defined a priori, were at significantly increased risk of pregnancy complications (OR = 5.0, 95% CI = 1.9, 12.6). The woman's own perception of her neighborhood milieu was also associated with an increase in risk (OR = 1.8, 95% CI = 0.8, 4.3 at interview 1 and OR = 1.8, 95% CI = 0.9, 3.6 at interview 2). The presence of violence in the home was not significantly associated with an increased risk of pregnancy complications (OR = 1.5, 95% CI = 0.8, 3.2).

Multiple logistic regression was used to assess the independent effects of age, low social support, high level of alienation, perception of neighborhood milieu, and sociopolitical violence in the neighborhood, that is, all variables for which $P < .20$ in previous analyses. Because several of the variables were not independently associated with the risk of pregnancy complications in the multivariate analysis, a second model included only the three variables that remained predictive of pregnancy complications at $P < .20$ (Table 5). A fivefold increase in the risk of pregnancy complications was associated with living in high-violence versus low-violence neighborhoods, after adjustments were made for lack of social support and perception of neighborhood milieu at interview 1.

Several variables were associated with living in high-violence neighborhoods, including income, education, marital status, underweight, cigarette smoking, dissatisfaction with neighborhood, life events (at interview 2), neighborhood milieu (at interview 1), alienation, uncertainty, and depression (at interview 1). Most of these variables were not directly associated with pregnancy complications. However, we chose to include them in a multiple logistic model as potential confounders of the association between living in a high-violence neighborhood and risk of pregnancy complications. The results of this multivariate analysis closely paralleled those shown in Table 5. The only predictive variables were sociopolitical violence of the neighborhood (OR = 5.3, $P < .01$), alienation (OR = 1.8, $P = .18$), and neighborhood milieu at interview 1 (OR = 1.9, $P = .17$). The *P* values for all other variables in the model were higher than .4.

To assess whether the apparent effect of neighborhood sociopolitical violence resulted from differences in diagnosis of pregnancy complications among the health centers, we tested for an association between specific health center and risk of pregnancy complications separately for high-violence neighborhoods and low-violence neighborhoods. The fact that no significant associations were ob-

served suggested that the levels of violence in the health centers' neighborhoods, rather than characteristics of the health centers themselves, influenced pregnancy complications. However, we did find an interaction between neighborhood violence and social support in these analyses. Within high-violence neighborhoods, lower levels of social support were significantly associated with risk of pregnancy complications (OR = 3.1, 95% CI = 1.3, 7.6). In other words, women exposed to high levels of violence in their neighborhoods who also had less social support were more likely to have pregnancy complications than were women in the high-violence neighborhoods whose social support was stronger.

Discussion

The purpose of this study was to determine whether there was a difference in pregnancy complications between women who live under high levels of sociopolitical violence and women who live under less violent conditions. After adjusting for potential confounders, we found that high levels of sociopolitical violence were associated with an approximately fivefold increase in risk of pregnancy complications. Two subjective measures, lack of social support and perceived sociopolitical violence (neighborhood milieu), also independently contributed approximately twofold increases in risk of pregnancy complications among otherwise healthy women. The latter results are similar to those found in other studies of pregnancy complications.[24–26]

One potential source of bias in this study is possible misclassification of women's exposure to violence. However, the use of objective indicators of sociopolitical violence to categorize neighborhoods, as well as the independent confirmation by the health centers' directors, gives us confidence in our mapping efforts. Moreover, the ability of the health centers' directors to identify pockets of high sociopolitical violence within areas classified as low-violence gave us the opportunity to identify more precisely the level of violence within neighborhoods and to assign women appropriately.

The prospective nature of the study precluded the possibility of assigning exposure status on the basis of outcome. Pregnancy complication information was obtained after completion of both interviews by medical record abstractors who were unaware of whether a woman resided in a high- or low-violence neighborhood. Moreover, the medical care providers did not know which pregnant women were included in our study. Nevertheless, we addressed the possibility that medical care or diagnostic practices may have differed by health center and thereby influenced our findings. The fact that these analyses revealed no association between health centers and pregnancy complications after we controlled for level of sociopolitical violence suggests that differential ascertainment of pregnancy complications was not a problem.

We were initially concerned about the respondents' willingness to participate fully in our prospective study. However, our response rate was 100% of those contacted and eligible at the first interview and 99% of those still living in the target areas and still pregnant at the second interview. The sample was relatively small, but it was representative of healthy pregnant women in Santiago who received prenatal care at the health centers. This group includes the majority of lower-middle-class, working class, and unemployed pregnant women in Santiago. Our conclusions cannot be generalized to pregnant teenagers, women with chronic illness, or those not receiving prenatal care.

Public Health Implications

Violence is a public health issue and deserves the attention accorded to any other cause of morbidity and mortality in human populations. Some countries have a specialization called trauma medicine, which deals with severe body damage caused by violence. New techniques in psychiatry and psychology have contributed to the healing of invisible wounds left by political violence. Two of the best examples of efforts to address violence are the centers for the rehabilitation of victims of torture in Denmark[27] and Los Angeles.[7]

Medical resources for curing or rehabilitating victims of violence are scarce, but preventive measures in the medical and public health disciplines are nonexistent. Most of the work on the prevention of violence in general, and political violence in particular, is being done by nonprofit groups and organizations. It is not too late for the public health field to effectively address its negligence of social and political violence and the threat such violence represents to the lives of individuals and communities.

To our knowledge, this is the first study of the effect of sociopolitical violence on pregnancy complications. Four women in our study miscarried after 20 weeks' gestation. All four resided in high-violence areas. Among the other women, complications were generally mild or moderate. The finding that healthy women were more subject to pregnancy complications if they lived in high-violence neighborhoods thus reflects unexpected health consequences of sociopolitical violence in a vulnerable population. The fact that 98% of these women carried their pregnancies to term despite the reality of violence in their neighborhoods during their pregnancies is a tribute, in our view, to their ability to cope and to the excellent care provided by their neighborhood health centers.

Research on health-related effects of social and political violence is very limited, particularly in communities that experience high levels of violence for prolonged periods of time. Clearly, this study requires confirmation with a larger sample and other populations. At the same time, data are needed on the long-term effects of sociopolitical violence on the postpartum and neonatal well-being of mothers and their infants, as well as on other sectors of the general population. As public health professionals, we can contribute to the control of social and political violence by actively engaging in the study and prevention of such violence. □

Acknowledgments

This study was supported by a Fulbright Foundation grant to (Dr Zapata, 1985–1986) and by the Grossman Fund of the University of California, Berkeley, School of Public Health. This research was approved by the Chilean Ministry of Health on February 18, 1986.

The results of this study were presented at the 117th Annual Meeting of the American Public Health Association, October 1989, in Chicago. We wish to thank the dean of the School of Medicine of the Universidad de Chile; Roberto Belmar; the directors and midwives in the six health centers; our five interviewers and medical records abstractors; Jeffrey Gould, Troy Duster, Thomas Boyce, Leonard Syme, the faculty at the Maternal and Child Health Program, David Ragland, Charles P. Quesenberry, Jr, Jerry DeLorenzo, and Robert Scholtz, all at the University of California, Berkeley; and Terry Camacho at the Human Population Laboratory in Berkeley.

References

1. The Centers for Disease Control Study of Vietnam Experience: post-service mortality among Vietnam veterans. *J Am Med Assoc.* 1987;257:790–795.
2. The Centers for Disease Control Study of Vietnam Experience: health status of Vietnam veterans: II. Physical health. *J Am Med Assoc.* 1988;259:2708–2714.
3. The Centers for Disease Control Study of Vietnam Experience: health status of Viet-

nam veterans: III. Reproductive outcomes and child health. *J Am Med Assoc.* 1988; 259:2715–2719.
4. Fett MJ, Nairn JR, Cobbin DM, Adena MA. Mortality among Australian conscripts of the Vietnam conflict era: II. Causes of death. *Am J Epidemiol.* 1987;125:878–884.
5. Welte JW, Abel EL. Homocide and race in Erie County, New York. *Am J Epidemiol.* 1986;124:666–670.
6. Paulson JA, Rushforth NB. Violent death in children in a metropolitan county: changing patterns of homicide, 1958 to 1982. *Pediatrics.* 1986;78:1013–1020.
7. Allobi F, Randall GR, Lutz EL, et al. Physical and psychiatric effects of torture: two medical studies. In: Stover E, Nightengale EO, eds. *The Breaking of Bodies and Minds: Torture, Psychiatric Abuses and the Health Profession.* New York: Freeman; 1985.
8. Daugaard G, Petersen HD, Abilgaard U, et al. Sequela to genital trauma in torture victims. *Arch Androl.* 1983;10:245–248.
9. Herman AA. Political violence, health, and health services in South Africa. *Am J Public Health.* 1988;8:767–768.
10. Yach D. The impact of political violence on health and health services in Capetown, South Africa, 1986: methodological problems and preliminary results. *Am J Public Health.* 1988;78:772–776.
11. Blumberg B, Corilion C, Debreu G, Stover E. *Scientists and Human Rights in Chile: Report of a Delegation.* Washington, DC: National Academy Press; 1985.
12. Morrison BA. Chile, the multilateral development banks and United States human rights laws: A delegation report. In: Hearing before the subcommittee on International Development Institutions and Finance of the Committee on Banking, Finance, and Urban Affairs, House of Representatives. Human Right Abuses in Chile: Time for United States Action (Serial No. 99-89). Washington, DC: United States Government Printing Office; 1986:73–112.
13. Brown C. Statement In: Hearing before the Subcommittee on International Development Institutes and Finance of the Committee on Banking, Finance, and Urban Affairs, House of Representatives. Current Directions for United States Policy toward Chile (Serial No. 99-89). Washington, DC: United States Government Printing Office; 1987:58–64.
14. Constable J. Statement In: Hearing before the Subcommittee on International Development Institutions and Finance of the Committee on Banking, Finance, and Urban Affairs, House of Representatives. Human Rights Abuses in Chile: Time for United States Action (Serial No. 99-89). Washington, DC: United States Government Printing Office; 1986:6–10.
15. Human Population Laboratory. *Psychosocial Scales: Indices of Social Support, Alienation, Depression, and Locus of Control.* Berkeley, Calif: California State Department of Health Services; 1987.
16. Eckenrode J, Gore S. Stressful events and social supports: the significance of context. In: Gottlieb BH, ed. *Social Network and Support in Community Mental Health.* Beverly Hills, Calif: Sage; 1981.
17. Berkman L, Breslow L. *Health and Ways of Living.* New York: Oxford University Press; 1983.
18. Eckenrode J. Impact of chronic and acute stressors on daily reports of mood. *J Pers Soc Psychol.* 1984;4:907–918.
19. Eckenrode J. The mobilization of social support: some individual constraints. *Am J Community Psychol.* 1983;11:509–528.
20. Zapata BC. *An Epidemiologic Study of the Influence of Social and Political Violence on the Risk of Pregnancy Complications.* Berkeley, Calif: University of California, Berkeley; 1989. Dissertation.
21. SAS Institute Inc. *SAS User's Guide: Basics, Version 5 Edition.* Cary, NC: SAS Institute; 1985.
22. SAS Institute Inc. *User's Guide: Statistics, Version 5 Edition.* Cary, NC: SAS Institute; 1985.
23. SAS Institute Inc. *SUGI Supplemental Library User's Guide, Version 5 Edition.* Cary, NC: SAS Institute; 1985:269–293.
24. Nuckolls KB, Cassal J, Kaplan BH. Psychosocial assets, life crisis, and the prognosis of pregnancy. *Am J Epidemiol.* 1972;95:431–441.
25. Norbeck JS, Tilden VP. Life stress, social support, and emotional disequilibrium in complications of pregnancy. *J Health Soc Behav.* 1983;34:30–46.
26. Norbeck JS, Anderson JN. Psychosocial predictors of pregnancy outcomes among low income black, hispanic, and white women. *Nurs Res.* 1989;38:204–209.
27. Somnier FE, Genefke IK. Psychotherapy for victims of torture. *Br J Psychiat.* 1986;149:323–329.

Sources of Fiber and Fat in Diets of US Women Aged 19 to 50: Implications for Nutrition Education and Policy

ABSTRACT

Introduction. Many women are not meeting dietary recommendations for fiber and fat intakes. Health educators need to know which foods are major sources of these nutrients.

Methods. This study analyzed the contributions of 27 food groups to fiber, fat, saturated fat, and cholesterol intakes of 2134 women aged 19 to 50 years in USDA's Continuing Survey of Food Intakes by Individuals, the 1985 and 1986 surveys (CSFII 85–86).

Results. Major determinants of fiber intake include frequency of use of certain food groups (vegetables, including potatoes, bread, fruit, soups, ready-to-eat cereal) and choice of particular foods within the larger food groups (e.g., whole grain bread, high fiber cereal). Major determinants of total fat, saturated fat, and cholesterol intakes included frequency of use of certain foods (sweet grains, beef, eggs, cheeses/cream, whole milks) and additions to foods (regular salad dressing and butter/margarine). Demographic characteristics were related to various food group consumption parameters.

Conclusions. Information about the relationships between food group and nutrient intake and the effects of household income, race/ethnicity, and region of residence on food group intake indicate opportunities to refine nutritional education programs. (*Am J Public Health*. 1992;82:695–702)

Frances E. Thompson, BS, MPH, PhD, Mary Fran Sowers, BS, MS, PhD, Edward A. Frongillo, Jr, ScB, MS, PhD, and Banoo J. Parpia, BA, BS, MA

Introduction

According to the Joint Nutrition Monitoring Evaluation Committee, the federal advisory committee jointly established by the US Departments of Agriculture (USDA) and Health and Human Services (HHS) to monitor the nutritional status of the US population, the principal nutrition-related health problems in the United States today arise from overconsumption of fat, saturated fatty acids, cholesterol, and sodium.[1] Epidemiological and clinical studies suggest that diets high in plant foods are associated with lower risk of coronary heart disease, colon cancer, diabetes mellitus, diverticulosis, hypertension, and gallstone formation.[2] In response to growing evidence concerning the relationships between diet and chronic diseases, a number of private and public organizations,[3–9] most recently USDA and HHS in their 1990 *Dietary Guidelines for Americans*,[10] have made specific quantitative recommendations about how much fiber and fat Americans should be consuming for optimal health.

Comparison of individual food consumption data collected by USDA in the 1977–78 Nationwide Food Consumption Survey (NFCS) and the 1985 Continuing Survey of Food Intakes by Individuals (CSFII 85) indicates that fat intake may have been decreasing in recent years. The mean intake of total fat for women 19 to 50 decreased from about 73.9 g per day[11] (40.8% of kilocalories from fat[2]) in 1977–78 to 63.1 g per day (36.8% of kilocalories from fat) in 1985.[12] Fiber intake was not estimated in the 1977–78 NFCS. The mean fiber intake for women 19 to 50, estimated from CSFII 85, was 10.6 g per day.[12]

Many women are not meeting dietary recommendations for fiber and fats. In order to provide effective advice to women on how to increase fiber and reduce fat intakes, health educators and nutritionists need to know which foods are major sources of these nutrients among various subgroups of women, defined by important demographic characteristics, and which eating practices account for differences in fiber and fat intakes. In this paper we address these questions.

Methods

Sampling and Study Procedures

Data from the USDA's Continuing Survey of Food Intakes by Individuals 1985 and 1986 (CSFII 85-86) core sample surveys were used. The design of these surveys provided two multistage stratified area probability samples of women aged 19 through 50 and their children aged 1 through 5, representative of the 48 conterminous states.

In order to ensure comparability of the CSFII 85 and the CSFII 86, the same area segments were sampled in the two surveys. However, different housing units were selected for each survey. The data collection and data processing procedures used in the two surveys were identical.

Survey participants were asked to provide six nonconsecutive 1-day dietary

The work was performed at the Division of Nutritional Sciences, Cornell University, Ithaca, NY.

Requests for reprints should be sent to Frances E. Thompson, Applied Research Branch, National Cancer Institute, Executive Plaza North, Room 330, Bethesda, MD 20892.

This paper was submitted to the Journal December 27, 1990, and accepted with revisions October 18, 1991.

TABLE 1—Description of 27 Food Groups Analyzed

Food Group	Foods Included
Beef	Beef, veal, lamb, and goat, solid and ground, except that contained in mixtures
Poultry	All poultry: chicken, turkey, and game birds, except that contained in mixtures
Fresh pork	Fresh (not cured) pork, except that contained in mixtures
Lunch meats, bacon	Franks, sausages, lunch meats, potted meat or spreads (including sandwiches), bacon, salt pork
Fish	All fish: finfish and seafood, except that contained in mixtures
Cured meat, organ meat, miscellaneous meat	Cured, smoked, pickled meat, game, organ meat, miscellaneous meats, and meat parts
Meat, fish, poultry mixtures	Beef, lamb, veal, goat mixtures, frozen plate meals, beef and veal sandwiches, pork mixtures, poultry mixtures, fish mixtures
Nuts, seeds	All nuts and seeds, nut butters
Eggs	Eggs and egg mixtures such as egg salad, omelettes
Whole milks	Whole milk, fluid and dry, flavored and plain; evaporated milk
Low-fat milks, cottage cheese, yogurt	Low-fat and skim milk, fluid and dry, low-calorie flavored milk drinks, yogurt, cottage cheese and other low-fat cheeses
Cheeses/cream	Natural cheese, processed cheese, cream cheese, cream and cream substitutes
Dairy/frozen desserts	Gelatin desserts, ices, popsicles, sherbert, puddings, ice cream
Soups/dried beans	All soups, including vegetable, meat, poultry, fish, and bean soups; dried beans
Mostly grain/vegetable mixtures	Mixtures with pasta or rice, pizza, Hispanic mixed dishes, Asian mixed dishes
Rice/pasta	Rice, pasta
Bread	Yeast breads, white and whole grain
Biscuit/cornbread/cooked cereals	Biscuits, grits, cornmeal mush, cornbread, cooked cereals
Ready-to-eat cereals	All ready-to-eat breakfast cereals
Fruits	All fruits, including dried fruits
Fruit juices	All fruit juices, nectars, tomato juice
White potatoes	White potatoes, prepared all ways
Vegetables, not white potatoes	All vegetables other than white potatoes and other than olives, pickles, relishes, and pickled vegetables
Sweet grain products	Sweet rolls, pancakes, pies, eclairs, cakes, cookies, sweet crackers
Salty snacks	Crackers and salty snacks, including popcorn and pretzels, potato chips/sticks
Butter/margarine	Butter, margarine
Regular salad dressing	Mayonnaise, all regular salad dressings

recalls collected at approximately 2-month intervals throughout a year. Interviews occurred from April 1985 through March 1987. Eligible households were scheduled for interviews on different days of the week to provide representativeness of dietary intake data over all days of the week. The first interview was conducted in person; subsequent interviews were conducted by telephone for 91% of the sample. Standard methods were used for conducting the interview.[12,13]

While the design of the survey called for collection of 6 nonconsecutive days of information, less than half of the CSFII 85-86 respondents interviewed initially completed the full 6 days of report. Consequently, USDA constructed a sample of all women who had completed any 4 of the possible 6 days of report, 71% of the initial CSFII 85 sample and 76% of the initial CSFII 86 sample. For all selected women, the first day of information was included. For those women who reported 5 or 6 days of information, USDA randomly selected the 3 additional days included. The combined CSFII 85-86 data set used in these analyses includes 2134 women, each with 4 days of dietary information. A more complete description of the sampling and study procedures is found in CSFII reports 85-4[12] and 86-3.[13]

Variable Definition

All individual foods reported were classified into 89 food groups based on patterns of use, frequency of use, and nutrient composition. Foods that were eaten together and reported as one food item, rather than reported as separate component foods, were categorized into mixture food groups (e.g., beef sandwiches). For each respondent, the total amount of the food group consumed in grams, the total frequency of food group use, and the amount of nutrient (e.g., fat) consumed from the food group were calculated for each food group. The 89 food groups were also aggregated into a system of 33 food groups. Most of these analyses used the system of 33 food groups, although some used the system of 89 food groups when a greater level of detail was required. Six of the 33 food groups contained no fat or fiber, and so were omitted from these analyses. The 27 food groups remaining are described in Table 1.

In order to examine the effects of specific eating practices on overall nutrient intake, the frequency of consuming each food group (measured as the number of times foods from the food group were consumed over the 4 days of report), the serving size (in grams) of the food group when it was consumed, and the nutrient density of the food group when it was consumed (e.g., the amount of fat present in a 100-g serving of that food group) were specified. The density measure reflects the fact that a different mix of foods with a different average nutrient composition might have been consumed by different individuals. For example, an individual who chose relatively high-fat hamburger would have a higher fat density in the ground beef food group than would an individual who chose lower fat hamburger.

Analytical Methods

In order to determine the major food group sources of fiber, fat, saturated fatty acids, and cholesterol in the population's diet, the mean amount of nutrient contributed by each food group to the population's total nutrient intake was estimated by averaging the nutrient from food group variable over all individuals. The mean proportion each food group contributed to nutrient intake was estimated as follows: The total amount of nutrient consumed from each food group was summed over all individuals. This was divided by the total amount of nutrient consumed from all food groups summed over all individuals. This analysis was performed on the total sample and on subgroups of the sample stratified by household income as a percentage of poverty ($\leq$185% vs $>$ 185%), ethnicity (White, Black, Hispanic; others omitted because of small numbers),

and region of residence (Northeast, Midwest, South, and West).

In order to compare specific eating practices of women at different levels of nutrient intake, the sample was divided into quartiles based on that nutrient intake. Mean levels of eating practice variables between women in the first quartile and women in the fourth quartile were compared.

Weighting factors that adjust for household and individual nonresponse were provided by USDA and incorporated in all analyses. Weighted proportions, quartile levels, and means adjusting for nonresponse were estimated using SAS.

Results

Intakes of Fiber and Fats for Demographic Subgroups

Mean daily intakes of fiber and fats for women aged 19 to 50 are shown in Table 2. White women have higher fat and saturated fatty acid intakes but lower cholesterol intakes and higher fiber intakes. Household income as a percentage of poverty is directly associated with fat, saturated fatty acids, and cholesterol intakes. Women in the Northeast have the lowest intakes of fat, saturated fatty acids, and fiber. Women in the Midwest and West have greater intakes of fiber and saturated fatty acids.

While the diets of many women are in accordance with the cholesterol recommendation, the diets of most women contain lower than recommended amounts of fiber and higher than recommended amounts of total fat and saturated fatty acids (Table 3). More Hispanics are meeting recommendations for fat and saturated fatty acids. Fewer Whites meet the saturated fatty acid recommendation, and fewer Blacks meet the cholesterol recommendation. Higher household income as a percentage of poverty is associated with lower prevalences of meeting the percentage of kilocalories from fat and saturated fatty acids recommendations and is unrelated to the prevalence of meeting the cholesterol recommendation.

TABLE 2—Weighted Mean Daily Nutrient Intake of Women 19–50 by Demographic Characteristics: CSFII 85-86

Demographic Characteristics	n	Fat, g	Kilocalories from Fat, %	Saturated Fat, g	Kilocalories from Saturated Fat, %	Cholesterol, mg	Fiber, g
Region							
Northeast	464	57.6	35.5	21.2	12.9	280	9.8
Midwest	583	66.0	37.3	24.4	13.7	276	11.2
South	685	60.0	35.8	21.6	12.8	270	10.4
West	402	64.2	37.0	23.5	13.5	279	11.8
Income as % of poverty							
≤130	455	58.6	36.2	21.4	13.1	283	9.3
130.1–300	815	60.1	35.7	21.8	12.8	271	10.7
>300	864	65.0	37.0	23.9	13.6	276	11.6
Race/ethnicity							
White	1773	63.4	36.7	23.3	13.4	271	11.1
Black	173	55.3	36.0	19.3	12.5	313	8.5
Hispanic	129	53.5	33.9	19.5	12.3	284	9.2
Total	2134	61.8	36.4	22.6	13.2	276	10.8

Note. CSFII 85-86 = Continuing Survey of Food Intakes by Individuals 1985 and 1986.

TABLE 3—Weighted Percentage of Women 19–50 Meeting Dietary Recommendations by Demographic Characteristics: CSFII 85-86

Demographic Characteristic	Dietary Component (recommendation)			
	Fat, (≤30%)[a,b]	Saturated Fat, (≤10%)[a,b]	Cholesterol, (≤300 mg/day)[a]	Fiber, (≥20 g/day)[c]
Region				
Northeast	15.4	15.8	60.6	4.2
Midwest	11.1	8.3	64.7	5.4
South	17.4	15.6	65.0	4.2
West	12.9	14.7	60.2	4.6
Income as % of poverty				
≤130	17.6	15.2	63.4	3.2
130.1–300	15.9	16.2	63.2	4.1
>300	11.5	10.7	62.4	5.7
Race/ethnicity				
White	13.3	11.5	64.9	4.9
Black	14.3	19.3	49.3	3.6
Hispanic	21.4	24.7	57.7	2.1
Total	14.5	13.7	62.9	4.6

Note. CSFII 85-86 = Continuing Survey of Food Intakes by Individuals 1985 and 1986.
[a]National Academy of Sciences.[2]
[b]US Departments of Agriculture and Health and Human Services.[10]
[c]National Cancer Institute.[15]

Fiber

Food group sources by poverty level. The food group vegetables other than potatoes contributes nearly a quarter of the total amount of fiber consumed in both lower and higher income women (Table 4). (Note that this does not include vegetables eaten in meat and grain mixtures or soups.) Fruits and ready-to-eat cereals are more important contributors to fiber intake among higher income women; white potatoes and soups/dried beans are more important contributors among lower income women.

Eating practices. Diets in the fourth (highest) quartile of fiber consumption average about 17.5 g of fiber daily; those in the first (lowest) quartile average about 5.4 g daily. Although women in the fourth quartile of fiber consumption generally consume the major fiber food groups with greater frequency, serving size, and fiber density than do women in the first quartile, the largest differences are in the frequency of eating component (data not shown). The mean frequency of use for six of the eight food groups examined (i.e., vegetables other than potatoes, bread, fruits, white potatoes, soups/dried beans, and ready-to-eat cereals but not grain mixtures and meat/fish/poultry mixtures) is substantially higher among women in the

TABLE 4—Weighted Food Group Sources of Fiber Intake among Women 19–50 by Household Income: CSFII 85-86

Household Income/Food Group	Fiber, g/day	% of Total	Household Income/Food Group	Fiber, g/day	% of Total
≤185% of poverty (n = 726)[a]			>185% of poverty (n = 1408)[b]		
Vegetables other than white potatoes	2.3	23.5	Vegetables other than white potatoes	2.8	24.4
Bread	1.2	12.5	Bread	1.5	13.3
White potatoes	1.1	11.1	Fruits	1.4	12.2
Soups	0.9	9.3	White potatoes	1.0	8.6
Fruits	0.8	8.3	Mostly grain, vegetable mixtures	0.7	6.6
Mostly grain, vegetable mixtures	0.7	7.2	Ready-to-eat cereals	0.6	5.7
Meat, fish, poultry mixtures	0.6	6.2	Meat, fish, poultry mixtures	0.6	5.2
Sweet grain products	0.4	3.9	Soups	0.6	5.1

Note. CSFII 85-86 = Continuing Survey of Food Intakes by Individuals 1985 and 1986.
[a]Weighted mean daily fiber intake is 9.7 g.
[b]Weighted mean daily fiber intake is 11.3 g.

fourth vs first fiber quartile. The mean serving size is somewhat greater for white potatoes, vegetables other than white potatoes, ready-to-eat cereals, grain mixtures, and meat/fish/poultry mixtures. The mean fiber density is substantially greater for ready-to-eat cereals and bread but not for the other six food groups examined.

We also examined the contribution of raw vegetables (mainly salads) to fiber differences between the two quartiles of women. Those in the highest vs lowest fiber quartile consume raw vegetables more frequently (2.7 vs 1.5 times over 4 days). Mean serving size and fiber density are similar.

Food patterns. In order to increase fiber, the USDA–HHS dietary guidelines,[10] the National Cancer Institute (NCI),[8] and the National Academy of Sciences[2] recommend that adults eat daily at least five servings of vegetables and fruits and six servings of grain products, several of which should be whole grain. Although women obtain fiber from these food groups, their frequency of consumption of these foods is much lower than advised. We estimate that among women 19 to 50, the average daily number of servings of vegetables (including legumes) is 2.6; of fruits or fruit juices, 0.8; and of grain products, 2.8 (Table 5). Only 10% eat a cereal that contains at least 5 g of fiber per serving at least once over 4 days. Furthermore, while 47% eat whole grain bread at least once over 4 days, only 26% of all bread consumed is whole grain.

The number of servings of vegetables, fruits, and grains consumed increases with increasing household income (Table 5), although consumption of potatoes and dried beans/soups is similar across three income groups. Whites consume the greatest number of servings of vegetables, fruits, and grains, followed by Hispanics and Blacks. Women in the West have the highest number of servings of total vegetables, total fruits/juices, and whole grain bread.

Total Fat

Food group sources by poverty level. Since fat is widely dispersed in our foods, nearly all of the 27 food groups contribute some fat to the population's total fat intake. Grain mixtures, meat/fish/poultry mixtures, beef, sweet grains, cheeses/cream, and butter/margarine are major sources of fat for both income groups (Table 6). However, lunch meats/bacon and whole milks are more important food group sources among lower income women, while regular salad dressing is more important among higher income women.

TABLE 5—Weighted Mean Daily Number of Servings of Food Groups among Women 19–50 by Demographic Characteristics: CSFII 85-86

Demographic Characteristics	n	Total Vegetables (≥3)[a]	Raw Vegetables	Total Fruits/Juices (≥2)[a]	Solid Fruits	Total Grains (≥6)[a]	Whole Grain Bread	High-Fiber Cereal
Region								
Northeast	464	2.5	0.6	0.9	0.5	2.5	0.21	0.03
Midwest	583	2.6	0.5	0.8	0.5	2.8	0.24	0.03
South	685	2.5	0.5	0.7	0.5	2.7	0.23	0.03
West	402	2.9	0.6	1.0	0.7	2.8	0.32	0.04
Income as % of poverty								
≤130	455	2.1	0.3	0.5	0.3	2.4	0.13	0.01
130.1–300	815	2.6	0.5	0.8	0.5	2.8	0.26	0.03
>300	864	2.9	0.6	1.0	0.7	2.8	0.32	0.04
Race/ethnicity								
White	1773	2.7	0.6	0.9	0.6	2.8	0.27	0.04
Black	173	2.1	0.4	0.6	0.3	2.4	0.14	0.01
Hispanic	129	2.4	0.4	0.8	0.4	2.5	0.19	0.01
Total	2134	2.6	0.5	0.8	0.5	2.8	0.25	0.03

Note. CSFII 85-86 = Continuing Survey of Food Intakes by Individuals 1985 and 1986.
[a]Standards from the US Department of Agriculture.[14]

TABLE 6—Weighted Food Group Sources of Fat Intake among Women 19–50 by Household Income: CSFII 85-86

Household Income/Food Group	Fat, g/day	% of Total	Household Income/Food Group	Fat, g/day	% of Total
≤185% of poverty (n = 726)[a]			>185% of poverty (n = 1408)[b]		
Beef	5.0	8.6	Regular salad dressing	5.2	8.3
Mostly grain, vegetable mixtures	4.9	8.3	Mostly grain, vegetable mixtures	5.1	8.0
Meat, fish, poultry mixtures	4.4	7.6	Sweet grain products	5.0	7.8
Lunch meats, bacon	4.3	7.3	Meat, fish, poultry mixtures	4.9	7.8
Sweet grain products	3.8	6.5	Beef	4.6	7.2
Natural, processed cheeses, cream	3.2	5.5	Natural, processed cheeses, cream	4.2	6.7
Butter/margarine	3.1	5.2	Butter/margarine	4.0	6.3
Whole milks	2.6	4.5	Lunch meats, bacon	3.6	5.7

Note. CSFII 85-86 = Continuing Survey of Food Intakes by Individuals 1985 and 1986.
[a]Weighted mean daily fat intake is 58.6 g.
[b]Weighted mean daily fat intake is 63.5 g.

Food group sources by region and race/ethnicity. Grain mixtures and beef are important sources of fat among all race/ethnicity groups (Table 7). Salad dressing is more important among Whites and among women in the West and Northeast; franks/bacon is more important among Blacks and among women in the Midwest and South; poultry is more important and meat mixtures and sweet grains are less important among Blacks; and grain mixtures are more important among Hispanics.

Eating practices. Dietary guidelines suggest using low-fat preparation methods, yet it appears that certain high-fat preparation methods are prevalent among women aged 19 to 50. For example, we estimate that about 71% of the time, poultry is eaten fried or breaded and/or with the skin not removed. Pork is eaten untrimmed and fried or breaded about 68% of the time.

We estimate that women in the fourth vs first quartile of fat consumption average 95.3 vs 33.6 g of fat daily, 40.4% vs 31.7% of kilocalories from fat. The mean frequency of use and the mean serving size when used are lower in women in the first vs fourth fat quartile for grain mixtures, sweet grains, regular salad dressing, lunch meats/bacon, butter/margarine, and cheeses/cream (data not shown). In addition, women in the fourth fat quartile choose foods within the food groups of meat mixtures and beef that have a somewhat greater amount of fat per 100-g serving than foods chosen by women in the first fat quartile.

Saturated Fatty Acids

Food group sources by poverty level. Fewer foods are sources of saturated fatty acids, since saturated fatty acids occur predominantly in animal foods and infrequently in plant foods. Grain mixtures, meat/fish/poultry mixtures, and sweet grains are important food group sources for all women (Table 8). Beef, whole milks, and lunch meats/bacon are more important contributors to the saturated fatty acid intake among lower income women, while cheeses/cream and butter/margarine are more important among higher income women.

Food group sources by region and race/ethnicity. Cheeses/cream, grain mixtures, and beef are important sources of saturated fatty acids among women in all regions (Table 9). Whole milks are more important among women in the Northeast and among Hispanics. Franks/bacon is more important among women in the Midwest and South and among Blacks.

Eating practices. Among women aged 19 to 50, we estimate that those in the fourth vs first quartile of saturated fatty acid consumption consume an average of

TABLE 7—Top Five Food Group Sources of Fat among Women 19–50 by Region and Race/Ethnicity: CSFII 85-86

Region								Race/Ethnicity					
Northeast (n = 449)		Midwest (n = 565)		South (n = 659)		West (n = 360)		White (n = 1773)		Black (n = 173)		Hispanic (n = 129)	
Food Group	%[a]	Food Group	%[a]	Food Group	%[a]	Food Group	%[a]	Food Group	%[a]	Food Group	%[a]	Food Group	%[a]
Beef	8.4	Grain mixtures	8.4	Beef	8.2	Grain mixtures	9.1	Grain mixtures	8.0	Franks/ bacon	8.8	Meat mixtures	8.5
Salad dressing	8.1	Meat mixtures	7.8	Meat mixtures	8.2	Sweet grains	8.3	Meat mixtures	7.7	Grain mixtures	7.8	Grain mixtures	7.8
Grain mixtures	8.0	Sweet grains	7.4	Grain mixtures	7.3	Salad dressing	8.3	Beef	7.6	Beef	8.2	Beef	7.8
Meat mixtures	7.5	Franks/ bacon	7.4	Sweet grains	7.2	Meat mixtures	7.1	Sweet grains	7.6	Poultry	7.4	Sweet grains	7.6
Cheeses/ cream	6.7	Butter/ margarine	7.1	Franks/ bacon	6.8	Beef	6.9	Salad dressing	7.5	Eggs	6.2	Eggs	6.1

Note. CSFII 85-86 = Continuing Survey of Food Intakes by Individuals 1985 and 1986.
[a]Proportion of fat provided by food group.

TABLE 8—Weighted Food Group Sources of Saturated Fatty Acid Intake among Women 19–50 by Household Income: CSFII 85-86

Household Income/Food Group	Saturated Fatty Acids, g/day	% of Total	Household Income/Food Group	Saturated Fatty Acids, g/day	% of Total
≤185% of poverty (n = 726)[a]			>185% of poverty (n = 1408)[b]		
Beef	2.1	9.7	Natural, processed cheeses, cream	2.7	11.6
Natural, processed cheeses, cream	2.0	9.5	Mostly grain, vegetable mixtures	1.9	8.2
Mostly grain, vegetable mixtures	1.9	8.8	Beef	1.9	8.2
Whole milks	1.6	7.6	Meat, fish, poultry mixtures	1.7	7.4
Meat, fish, poultry mixtures	1.6	7.4	Butter/margarine	1.5	6.5
Lunch meats, bacon	1.6	7.3	Sweet grain products	1.5	6.5
Sweet grain products	1.2	5.4	Lunch meats, bacon	1.3	5.7
Butter/margarine	1.1	4.9	Whole milks	1.3	5.6

Note. CSFII 85-86 = Continuing Survey of Food Intakes by Individuals 1985 and 1986.
[a]Weighted mean daily saturated fat intake is 58.6 g.
[b]Weighted mean daily saturated fat intake is 63.5 g.

35.9 vs 11.5 g of saturated fatty acids daily, 15.6% vs 10.7% of kilocalories from saturated fatty acids.

The mean frequency of use and mean serving size when used are substantially lower among women in the first vs fourth quartile of saturated fatty acid intake for cheeses/cream, grain mixtures, whole milks, lunch meats/bacon, butter/margarine, and sweet grains (data not shown). In addition, women in the fourth quartile vs first quartile are consuming butter/margarine of higher saturated fatty acid density.

Cholesterol

Food group sources by poverty level. The food group eggs contributes about a third of the cholesterol intake among women in lower and higher income groups (Table 10). Beef, meat/fish/poultry mixtures, grain mixtures, poultry, and sweet grains are also major contributors among women in both income groups. Organ, cured, miscellaneous meats and whole milks are more important contributors among lower income women. Cheeses/cream and fish are more important contributors among higher income women.

Food group sources by region and race/ethnicity. Differences in food group sources of cholesterol are less striking, since eggs constitutes the dominant food group source across all demographic subgroups (data not shown). Beef and meat mixtures are also important sources across all regions. Poultry and cured, organ, miscellaneous meats are more important among Blacks. Sweet grains is more important among women in the West.

Eating practices. Among women aged 19 to 50, we estimate that those with diets in the fourth vs first quartiles of cholesterol consumption average 472 vs 124 mg of cholesterol daily. The mean frequency of use and the mean serving size when used are lower among women in the first vs fourth quartile of cholesterol intake for eggs, beef, grain mixtures, and sweet grains (data not shown). These differences are especially substantial for eggs (0.2 vs 2.1 times in 4 days; 40- vs 87-g serving size). The cholesterol density of grain mixtures and sweet grains is higher among women in the fourth vs first cholesterol quartile.

Discussion

Major sources of fiber for women aged 19 to 50 are not only the foods in vegetables, grains, and fruits food groups eaten by themselves, but also in mixtures

TABLE 9—Top Five Food Group Sources of Saturated Fatty Acids among Women 19–50 by Region and Race/Ethnicity: CSFII 85-86

Region								Race/Ethnicity					
Northeast (n = 449)		Midwest (n = 565)		South (n = 659)		West (n = 360)		White (n = 1773)		Black (n = 173)		Hispanic (n = 129)	
Food Group	%[a]	Food Group	%[a]	Food Group	%[a]	Food Group	%[a]	Food Group	%[a]	Food Group	%[a]	Food Group	%[a]
Cheeses/cream	11.5	Cheeses/cream	9.6	Cheeses/cream	11.0	Cheeses/cream	12.0	Cheeses/cream	11.4	Beef	9.7	Whole milks	10.5
Beef	9.4	Grain mixtures	9.1	Beef	9.4	Grain mixtures	8.9	Beef	8.5	Franks/bacon	9.3	Cheeses/cream	10.4
Whole milks	8.2	Beef	7.9	Meat mixtures	8.1	Beef	7.9	Grain mixtures	8.3	Grain mixtures	9.1	Beef	8.8
Grain mixtures	8.0	Meat mixtures	7.7	Grain mixtures	7.6	Sweet grains	7.1	Meat mixtures	7.4	Cheeses/cream	6.9	Meat mixtures	8.5
Butter/margarine	6.9	Franks/bacon	7.4	Franks/bacon	6.9	Meat mixtures	6.8	Butter/margarine	6.4	Meat mixtures	6.0	Sweet grains	7.4

Note. CSFII 85-86 = Continuing Survey of Food Intakes by Individuals 1985 and 1986.
[a]Proportion of saturated fatty acids provided by food group.

TABLE 10—Weighted Food Group Sources of Cholesterol Intake among Women 19–50 by Household Income: CSFII 85-86

Household Income/Food Group	Cholesterol, mg/day	% of Total	Household Income/Food Group	Cholesterol, mg/day	% of Total
≤185% of poverty (n = 726)[a]			>185% of poverty (n = 1408)[b]		
Eggs	99	35.5	Eggs	83	30.4
Beef	27	9.6	Beef	26	9.6
Meat, fish, poultry mixtures	21	7.4	Meat, fish, poultry mixtures	24	8.6
Mostly grain, vegetable mixtures	18	6.5	Sweet grain products	19	6.9
Poultry	18	6.3	Mostly grain, vegetable mixtures	18	6.5
Sweet grain products	14	5.1	Poultry	18	6.4
Organ, cured, miscellaneous meats	12	4.1	Natural, processed cheeses, cream	13	4.6
Whole milks	11	3.8	Fish	11	4.1

Note. CSFII 85-86 = Continuing Survey of Food Intakes by Individuals 1985 and 1986.
[a]Weighted mean daily cholesterol intake is 280 mg.
[b]Weighted mean daily cholesterol intake is 273 mg.

(i.e., grain mixtures, soups, and meat/fish/poultry mixtures). The frequency of consumption of these food groups and the choice of particular foods within the food groups are creating large differences in fiber intake.

Women with the highest fiber intakes consume vegetables (including raw vegetables), bread, and fruits much more frequently than do those with the lowest fiber intakes. Women with higher fiber diets are also choosing higher fiber kinds of bread and ready-to-eat cereals.

However, few women are meeting the NCI recommendation of 20 g daily of fiber. In order to meet this recommendation using dietary guidance provided by USDA, and NCI[14,15] women need to increase their vegetable consumption and double their consumption of grain products and fruits.

For fats, it appears that the frequency of consumption of certain food groups, the use of certain additions to foods, and preparation methods used for these foods are responsible for the large differences in fat intake among women aged 19 to 50.

Disappearance data indicate that although there has been a decrease in visible separable fat consumption, there has also been an increase in "hidden fat" consumption through processed foods (e.g., baked goods), fried foods, and foods consumed in restaurants.[16] About 50% of the fats and oils in the food supply are used in processed foods.[16] Restaurant use of fats and oils has increased dramatically, primarily due to an increase in the number of restaurants serving fried foods.[16] As red meat consumption decreases, more meat is being purchased as nonseparable fat (e.g., ground beef, hotdogs) than separable fat.[16] Consumers may be unaware of the levels of fat in these products, since current grading standards are detrimental to consumer understanding of the fat content of meats.[16] Our results for individuals are consistent with these aggregated data; for example, franks/bacon and sweet grain products were major contributors to total intakes of fat, saturated fatty acids, and cholesterol. The challenge for nutrition educators is to help consumers identify these sources of hidden fats, both at point of purchase and in restaurants.

The use of high-fat additions appears to be an important determinant of total fat and saturated fatty acid intakes. Results indicate that women with the highest fat intakes consume regular salad dressing and butter/margarine more frequently and at greater serving sizes than do women with the lowest fat intakes. Women with the highest saturated fatty acid intakes consume butter/margarine more frequently, at greater serving sizes, and choose higher saturated fatty acid types more frequently, than do women with the lowest saturated fatty acid intakes. Nutrition education messages must emphasize that increasing consumption of salads, vegetables, and grain products as a means of reducing fat intake is effective only if high-fat additions are used sparingly.

Our results indicate that nutrition education messages about preparation methods such as breading, frying, and removing the skin of poultry would be beneficial to women aged 19 to 50. Among this population, it appears that meat/fish/poultry mixtures and grain mixtures are major sources of both fiber and total fat, saturated fatty acid, and cholesterol intakes, possibly due to sauces, butter/margarine, or cheese added to these foods in preparation. Messages about low-fat food preparation methods (e.g., use of low-fat rather than higher fat sauces and cheeses) should be devised specifically for these food groups.

Differences in the relative contributions of various food groups to fat intakes among different demographic subgroups can be used to refine educational messages. For example, substitution of low-fat for higher fat dairy products is more important for Hispanics and Whites; substitution of lower fat foods for franks/bacon is more important for Blacks. □

Acknowledgments

This study was carried out through the support of the Human Nutrition Information Service, US Department of Agriculture, Cooperative Agreement 58-3198-6-34. The research was conducted using the Cornell National Supercomputer Facility, a resource of the Cornell Theory Center, which is funded in part by the National Science Foundation, New York State, the IBM Corporation, and members of the Center's Corporate Research Institute.

Portions of this paper were presented at the 65th Agricultural Outlook Conference, Washington, DC, November 1988.

References

1. US Department of Agriculture/US Department of Health and Human Services. *Nutrition Monitoring in the United States: An Update Report on Nutrition Monitoring.* Washington, DC: US Government Printing Office; 1989. DHHS publication no. (PHS) 89-1255.
2. National Academy of Sciences, National Research Council Committee on Diet and Health. *Diet and Health. Implications for Reducing Chronic Disease Risk.* Washington, DC: National Academy Press; 1989.
3. American Heart Association, Commission on Nutrition. Diet and coronary heart disease. *Nutr Today.* 1974;9:26.
4. American Heart Association. Dietary guidelines for healthy adult Americans. *Circulation.* 1986;74:1465A.
5. National Institutes of Health. *NIH Consensus Development Statement on Lower-*

ing Blood Cholesterol to Prevent Heart Disease. Washington, DC: National Institutes of Health; 1984;5(7).
6. Committee on Diet, Nutrition, and Cancer, Assembly of Life Sciences, National Research Council. *Diet, Nutrition, and Cancer.* Washington, DC: National Academy Press; 1982.
7. American Cancer Society. Nutrition and cancer, cause and prevention. An American Cancer Society special report. *CA.* 1984;34:121–126.
8. Butrum RR, Clifford CK, Lanza E. NCI dietary guidelines: rationale. *Am J Clin Nutr.* 1988;48(suppl).
9. US Department of Health and Human Services. *The Surgeon General's Report on Nutrition and Health 1988.* Washington, DC: US Government Printing Office; 1988. DHHS publication no. (PHS) 88-50210.
10. US Department of Agriculture and US Department of Health and Human Services. *Nutrition and Your Health: Dietary Guidelines for Americans.* 3rd ed. Washington, DC: US Government Printing Office; 1990.
11. Larkin FA, Guire K, Hinton P, Metzner HL, Thompson FE. *The Comparison of Data from the Nationwide Food Consumption Survey and the National Health and Nutrition Examination Survey I. Final Report to the Human Nutrition Information Service, U.S. Department of Agriculture, 1982; University of Michigan.*
12. US Department of Agriculture, Human Nutrition Information Service. *Nationwide Food Consumption Survey Continuing Survey of Food Intakes by Individuals, Women 19–50 Years and Their Children 1–5 Years, 4 Days, 1985.* Hyattsville, MD; 1987. CSFII Report No. 85-4.
13. US Department of Agriculture, Human Nutrition Information Service. *Nationwide Food Consumption Survey Continuing Survey of Food Intakes by Individuals, Women 19–50 Years and Their Children 1–5 Years, 4 Days, 1986.* Hyattsville, MD; 1988. CSFII Report No. 86-3.
14. US Department of Agriculture, Human Nutrition Information Service. *Dietary Guidelines and Your Diet.* Washington, DC: US Government Printing Office; 1986. Home and Garden Bulletins 232-1–232-7.
15. National Cancer Institute. *Diet, Nutrition, and Cancer Prevention: A Guide to Food Choices.* Washington, DC: US Government Printing Office; 1987; NIH publication no. 87-2878.
16. National Research Council Committee on Technological Options to Improve the Nutritional Attributes of Animal Products. *Designing Foods: Animal Product Options in the Marketplace.* Washington, DC: National Academy Press; 1988.

1992 World Congress on Tuberculosis

The Fogarty International Center of the National Institutes of Health and the Centers for Disease Control, in collaboration with the National Institute of Allergy and Infectious Diseases, the National Heart, Lung and Blood Institute, the Food and Drug Administration, the US Agency for International Development, the World Bank, the World Health Organization, and the International Union Against Tuberculosis and Lung Disease, are pleased to announce that a World Congress on Tuberculosis will be held on November 16 to 19, 1992, in Bethesda, Maryland.

Invited to participate in the Congress are research scientists, tuberculosis control program officers, tuberculosis health care providers, and health officials, including Ministers of Health.

The Congress will focus on the goal of global tuberculosis control through basic and applied research in the diagnosis, treatment, and prevention of tuberculosis as well as behavioral and operational research. Program evaluation, quality assurance, and the epidemiologic, economic, and social aspects of tuberculosis will be discussed. Problems of special significance, such as the impact of HIV and drug resistance, will be addressed.

The Congress will feature addresses by leading national and international health officials; scientific presentations by leading investigators; and scientific poster sessions dealing with topics such as immunology, molecular biology, mycobacteriology, epidemiology, diagnosis, vaccine development, new drugs and therapy regimens, behavioral research, operational research, economic analyses, demonstrations/pilot projects, tuberculosis control programs, and program evaluation.

For additional information, or to obtain an official abstract form for presenting a poster, please contact the 1992 World Congress on Tuberculosis, Fogarty International Center, National Institutes of Health, Building 31, Room B2C32, Bethesda, Maryland 20892. Telephone number: 301/496-2516; Fax number: 301/402-2056.

Overcoming the Absence of Socioeconomic Data in Medical Records: Validation and Application of a Census-Based Methodology

Nancy Krieger, PhD

ABSTRACT

Background. Most US medical records lack socioeconomic data, hindering studies of social gradients in health and ascertainment of whether study samples are representative of the general population. This study assessed the validity of a census-based approach in addressing these problems.

Methods. Socioeconomic data from 1980 census tracts and block groups were matched to the 1985 membership records of a large prepaid health plan (n = 1.9 million), with the link provided by each individual's residential address. Among a subset of 14 420 Black and White members, comparisons were made of the association of individual, census tract, and census block-group socioeconomic measures with hypertension, height, smoking, and reproductive history.

Results. Census-level and individual-level socioeconomic measures were similarly associated with the selected health outcomes. Census data permitted assessing response bias due to missing individual-level socioeconomic data and also contextual effects involving the interaction of individual- and neighborhood-level socioeconomic traits. On the basis of block-group characteristics, health plan members generally were representative of the total population; persons in impoverished neighborhoods, however, were underrepresented.

Conclusions. This census-based methodology offers a valid and useful approach to overcoming the absence of socioeconomic data in most US medical records. (*Am J Public Health.* 1992;92:703–710)

Introduction

One problem frequently encountered by public health researchers and health planners in the United States is the absence of socioeconomic data in many widely used and routinely collected sources of health and disease information, such as medical records, disease registries, and vital statistics.[1,2] This omission is important, because social class, education, and income are known to influence both health status and use of health services.[1–7] Without these data, it is difficult not only to assess the role of socioeconomic factors in shaping the public's health, but also to ascertain whether a given study population is representative of the general population, and hence whether the study's findings legitimately can be generalized to the public at large.

To overcome this obstacle, various US investigators have sought to supplement individual-level health records with social class measures derived from an additional and easily available source of socioeconomic data, the US census.[1,8–20] In this approach, individuals are characterized by the socioeconomic profile of the immediate neighborhood in which they live (usually defined as the census tract), with the link provided by each person's residential address. To date, however, little research has examined the validity of this methodology,[1,8,21] including the degree to which it may result in ecologic fallacy (i.e., the erroneous inference of causal relations at the individual level on the basis of grouped data).[8,22–23]

To evaluate this census-based approach, I obtained health records with individual-level socioeconomic data from a large, multiethnic prepaid health plan and linked each individual's record to census-derived socioeconomic data acquired from two different geographic levels: (1) the census tract and (2) the census block group (a smaller and often more homogenous subdivision of the census tract that, on average, contains 1000 persons, as compared with the 4000 that typically reside in a census tract[24]). I then compared the association of both individual-level and census-based socioeconomic measures with four different health characteristics known to vary by race and socioeconomic position: hypertension,[2,25] height,[2,7,26] cigarette smoking,[2,7] and number of full-term pregnancies.[2,27] Census-based socioeconomic measures also were used to assess whether members of this health plan are representative of the surrounding general public. These analyses extend a smaller pilot project that compared individual, household, and census-derived social class measures as correlates of Black–White differences in reproductive history.[21]

Methods

Population

Study subjects consisted of all persons contained in the June 1985 computer-stored membership tape of the Kaiser Permanente Medical Care Program (KPMCP), Northern California Region (n = 1 924 995). Only data on age and gender are available for the total KPMCP membership; race/ethnicity and socioeco-

The author is with the Division of Research, Kaiser Foundation Research Institute, Oakland, Calif.

Requests for reprints should be sent to Nancy Krieger, PhD, Division of Research, Kaiser Foundation Research Institute, 3451 Piedmont Avenue, Oakland, CA 94611.

This paper was submitted to the Journal April 2, 1991, and accepted with revisions August 22, 1991.

nomic position are not ascertained. In 1987, the Medical Economics Department of KPMCP sent the June 1985 membership tape to be "geocoded" by a commercial firm, at a cost of $4.50 per 1000 records. This process assigned each individual a county code, census tract number, and block-group number based on the member's residential address and on geographic regions defined in the 1980 census.

A total of 1 593 388 KPMCP members (82.8%) were successfully geocoded. These members lived within 22 counties that were home to 98.8% of the total 1985 Northern California Region KPMCP membership. Missing or incorrect addresses, along with the absence of block-group codes in the 1980 census for three small Northern California counties (Amador, Lake, and Nevada), precluded fully geocoding the remainder. Members who were and were not successfully geocoded did not differ by gender and only differed slightly by age.

Census-Derived Data

Census tract and census block-group data from the 1980 US census were matched to each geocoded member's individual record. Each tract and block group was characterized by its social class and race/ethnic composition, and also by poverty and educational level.

Predominantly working-class tracts and block groups were defined as neighborhoods where 66% or more of the employed population belonged to the following census-based occupational categories: clerical and administrative support, sales, private household and other service occupations (except protective services), craft, transportation, and laborers. These occupations were selected because they disproportionately contain people who can be considered working class (i.e., employees who do not own their workplace, are not self-employed, and generally occupy subordinate positions at work).[21,28] The remaining census-defined occupations were considered non–working class: executive, administrative, and managerial; professional specialty; technicians and related support; protective service; and farming, forestry, and fishing (which includes farm owners and managers).

Impoverished tracts and block groups consisted of federally defined poverty areas, in which 20% or more of the population lived below the poverty line.[29] This economic indicator was chosen over the more commonly used, statistically defined measure of median family income because it takes into account a family's size and age structure and also pertains to the ability to buy a specified market basket of goods.[1,30,31] In 1980, the poverty level was set at $7356 for a family of two adults and two children.[29] Undereducated tracts and block groups were defined as regions where at least 25% of persons 25 years old or more had not completed high school.

Multiphasic Health Check-up Examination Data

Data from the multiphasic health check-up (MHC) examination were available for 17 200 persons contained in the geocoded June 1985 membership tape. This group represented 91.0% of the 18 904 KPMCP members who completed both the questionnaire and laboratory portions of the MHC examination in 1985 and who were KPMCP members as of June 1985.

Data obtained from the MHC examination included age at MHC examination, gender, race, years of completed education, occupation (current, if employed, or last occupation, if unemployed or retired), diastolic and systolic blood pressure, height, weight, cigarette-smoking history, and, for women, number of full-term pregnancies (NFTP). No data on income were collected. Using the same occupational groupings described above, social class was categorized as working class or non–working class. High blood pressure (HBP) was defined as diastolic ($\geq$ 90 mm Hg) and/or systolic ($\geq$ 145 mm Hg).[32] The Quetelet index (weight/height^2) served as a measure of body mass. The final sample consisted of 8674 White and 5566 Black KPMCP members included among the 17 055 persons in the MHC subset for whom race was identified.

Statistical Methods

Comparisons of demographic and health-related characteristics used the t test for continuous variables and the χ^2 test for categorical variables. The association of socioeconomic measures with the four health characteristics was determined by multivariate regression models run separately for the individual, tract, and block-group data. Analyses were performed with SAS programs for the personal computer.[33]

Multiple linear regression was used for health-related outcomes treated as continuous variables (height, NFTP). These models, in the form

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k,$$

permit estimations of how much the dependent variable (y) increases or decreases with a unit change in each independent variable (x_k), as indicated by its coefficient β_k, controlling for the other independent variables in the model.[33,34] Logistic regression was used for dichotomous (yes or no) outcomes (HBP, smoking status). This method compares the probability (p) of experiencing an outcome given a particular exposure x_k, and is expressed in the form:[33,35]

$$\text{logit}\,(p) = \log\,(p/(1 - p)) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k.$$

The relative risk associated with a unit change in each independent variable x_k, again controlling for other independent variables included in the model, can be calculated as $e^{\beta k}$.

For each health characteristic, the same group of individuals was analyzed at the individual, tract, and block-group level, with all results adjusted for the same relevant individual-level risk factors (e.g., age and gender). These analyses are referred to as the complete models, since they included only those individuals missing no individual-level or census-derived data for any of the variables included at any of the three model levels. Total models permitted analyzing the effect of response bias due to missing individual-level socioeconomic data, and included persons missing these data but not census-level socioeconomic data or other relevant individual-level data. Contextual regression analyses,[36,37] which incorporate individual- and group-level measures of the same traits (e.g., individual class position and working-class composition of that individual's tract or block group), were used to determine whether the effects of individual socioeconomic position are modified by neighborhood conditions. These models used the individual- and census-level data in the form of a combined dummy variable.

On the basis of the results of these analyses, block-group characteristics of the total geocoded KPMCP membership were compared with those of the total population for all 22 counties combined. Because of the extremely large size of the KPMCP sample (1.6 million, representing 21.6% of the total population), statistical tests were not used. Instead, differences between these two groups are noted and assessed.

TABLE 1—Comparison of Demographic and Health-Related Characteristics

	White		Black		Comparison of Blacks to Whites[b] (95% Confidence Interval)
	Value[a]	% missing data	Value[a]	% missing data	
Age, y (mean ± SD)	51.8 ± 16.3	0.0	47.8 ± 15.4	0.0	−4.0 (−5.3, −2.7)
Gender: women, %	54.2	0.0	60.3	0.0	1.11 (1.07, 1.14)
Class: working class, %	45.2	41.0	63.7	39.8	1.41 (1.36, 1.47)
Education: less than high school, %	22.7	24.9	38.6	25.6	1.70 (1.60, 1.80)
Height, cm (mean ± SD)	168.8 ± 9.9	0.1	168.1 ± 9.4	0.1	−0.7 (−1.0, −0.4)
Weight, kg (mean ± SD)	72.4 ± 15.4	0.1	78.1 ± 17.2	0.0	5.7 (5.1, 6.3)
Blood pressure, mm Hg					
Diastolic (mean ± SD)	75.3 ± 9.6	0.2	78.6 ± 10.4	0.2	3.3 (3.0, 3.6)
Systolic (mean ± SD)	130.7 ± 19.4	0.1	133.3 ± 19.1	0.1	2.6 (2.0, 3.3)
High Blood pressure[c]: yes, %	28.3	0.0	35.1	0.0	1.24 (1.18, 1.30)
Ever smoke: yes, %	51.7	24.9	53.1	25.2	1.03 (0.94, 1.07)
Still smokes: yes, %	33.8	4.1[d]	53.4	3.1[d]	1.58 (1.48, 1.68)
Full-term pregnancies (FTP) (mean ± SD)	1.6 ± 1.5	25.2[e]	2.1 ± 2.0	24.9[e]	0.5 (0.4, 0.6)
Census tract					
66+% working class, %	35.1	2.1	77.4	0.8	2.21 (2.14, 2.28)
25+% less than high school, %	20.4	2.8	67.2	0.8	3.29 (3.14, 3.44)
20+% less than poverty, %	10.2	2.1	43.0	0.8	4.22 (3.94, 4.53)
Census block group					
66+% working class, %	36.7	0.2	76.9	0.1	2.10 (2.03, 2.16)
25+% less than high school, %	24.5	0.2	66.6	0.1	2.72 (2.61, 2.83)
20+% less than poverty, %	11.1	0.2	43.3	0.1	3.90 (3.65, 4.17)

Note. 8674 White and 5566 Black KPMCP members.
[a]Value among those not missing data for the specified variable.
[b]Comparison: absolute difference (Black–White) for continuous variables, relative prevalence (Black ÷ White) for categorical variables.
[c]High blood pressure defined as diastolic (≥90 mm Hg) and/or systolic (≥145 mm Hg).
[d]Universe: all ever smokers (3371 Whites and 2210 Blacks).
[e]Universe: all women (4700 Whites and 3358 Blacks).

Results

Characteristics of the MHC Subset

The extent of missing data varied considerably for the demographic and health-related traits recorded at the MHC examination (Table 1). With the exception of the few persons missing tract-level data, Black and White MHC study subjects did not significantly differ in the proportion missing data for any given trait. They did differ significantly, however, for every demographic and health-related outcome except one (percentage ever smoked) (Table 1).

Validation of Census-Based Socioeconomic Measures

Among persons included in the complete models, measures of social class and educational attainment at the individual, census-tract, and block-group levels yielded similar estimates for the association of these socioeconomic variables with, respectively, HBP, height, still smoking (among ever smokers), and NFTP (Table 2). Block-group models generally yielded tighter confidence intervals than tract models. Tract and block-group models both tended to underestimate slightly the socioeconomic associations observed at the individual level, and neither rejected the null hypothesis (no class effects) when this hypothesis was not rejected by individual-level models. For each outcome, Black–White differences were similarly reduced by adjusting for individual- and census-level socioeconomic measures (Table 2).

Use of Census-Based Measures to Assess Respondent Bias

Adding persons who lacked individual- but not census-level socioeconomic data markedly increased the number of available subjects (Table 3). These additional persons tended to be older, and were more likely to be women and to reside in predominantly working-class, impoverished, and undereducated block groups, than persons included only in the complete models.

Evidence of respondent bias in the complete models was most apparent for Black–White comparisons concerning HBP and NFTP (Table 3). The adjusted relative risk of HBP among Black as compared with White MHC study subjects was 12% lower in the total than in the complete models (1.5 vs 1.7), while the relative excess of NFTP was 20% lower (0.4 vs 0.5). As would be expected, including substantially more persons in the total models produced tighter confidence intervals.

Contextual Analyses

Contextual regression analyses were conducted only with persons included in the complete models, since only these subjects possessed both individual- and census-level social class or educational measures. Suggestive contextual effects for both measures were observed for all four outcomes in both tract and block-group models. For example, as shown in Table 4, working-class women who lived in non–working-class block groups had virtually the same number of full-term pregnancies as non–working-class women who likewise lived in non–working-class block groups, whereas working-class women who lived in predominantly working-class block groups had 0.4 more full-term pregnancies than this latter group ($P < .05$).

Application of the Census-Based Methodology

A comparison of block-group characteristics of the geocoded 1985 KPMCP membership and the 1980 general population in the 22 corresponding counties re-

TABLE 2—Association of Race with Dependent Variables

	Model Characteristics					Risk Estimates[a] (95% Confidence Intervals)			
	Type of Regression Model	N White	N Black	Models Adjusted for	Model level	Race	Race +	Class +	Education
HBP[b]	Logistic	4960	3256	Age	Individual	1.9 (1.7, 2.1)	1.8 (1.6, 2.0)	1.0 (0.9, 1.2)	1.3 (1.2, 1.5)
				Gender	Tract	1.9 (1.7, 2.1)	1.7 (1.5, 1.9)	1.0 (0.9, 1.2)	1.2 (1.0, 1.4)
				Quetelet index	Block group	1.9 (1.7, 2.1)	1.7 (1.5, 1.9)	1.1 (0.9, 1.3)	1.2 (1.0, 1.4)
Height, cm	Linear	4975	3266	Age	Individual	−0.4 (−0.7, 0.1)	−0.1 (−0.4, 0.2)	−0.7 (−1.0, −0.4)	−1.1 (−1.4, −0.7)
				Gender	Tract	−0.4 (−0.7, −0.1)	−0.0 (−0.3, 0.4)	−0.4 (−0.9, −0.0)	−0.5 (−0.9, −0.0)
					Block group	−0.4 (−0.7, −0.1)	0.0 (−0.3, 0.4)	−0.7 (−1.0, 0.3)	−0.4 (−0.8, 0.0)
Smokes	Logistic	2544	1703	Age	Individual	2.1 (1.8, 2.4)	1.9 (1.7, 2.2)	1.4 (1.2, 1.6)	1.4 (1.2, 1.7)
				Gender	Tract	2.1 (1.8, 2.4)	1.8 (1.5, 2.1)	1.2 (1.0, 1.4)	1.2 (1.0, 1.5)
					Block group	2.1 (1.8, 2.4)	1.8 (1.6, 2.1)	1.2 (1.0, 1.4)	1.2 (1.0, 1.4)
NFTP	Linear	2535	1932	Age	Individual	0.7 (0.6, 0.8)	0.6 (0.5, 0.7)	0.1 (0.0, 0.2)	0.4 (0.3, 0.6)
					Tract	0.7 (0.6, 0.8)	0.5 (0.4, 0.7)	0.2 (0.0, 0.4)	0.1 (−0.1, 0.3)
					Block group	0.7 (0.6, 0.8)	0.5 (0.4, 0.7)	0.2 (0.0, 0.3)	0.2 (0.0, 0.3)

Note. HBP = high blood pressure; NFTP = number of full-term pregnancies.
[a]Comparison groups:
Race: Black vs White
Class: Individual = working class vs non-working class
Census tract = 66+% working-class vs <66% working-class census tract
Block group = 66+% working-class vs <66% working-class census block group
Education: Individual = <high school vs. ≥high school
Census tract = 25+% <high school vs <25% <high school census tract
Block group = 25+% <high school vs <25% <high school census block group
For linear regression: estimate of unit change in outcome associated with unit change in race, class, and education measures
For logistic regression: estimate of relative risk of outcome associated with unit change in race, class, and education measures
[b]HBP defined as diastolic (≥90 mm Hg) and/or syatolic (≥ 145 mm Hg).

vealed few differences with regard to race/ethnic composition (Table 5). Although a comparable proportion of KPMCP members and the general population lived in predominantly working-class block groups, KPMCP members were somewhat less likely than the general public to live in impoverished and undereducated block groups (Table 6).

Discussion

The census-based methodology of supplementing individuals' health records with socioeconomic data from their census block groups appears to offer a valid and useful approach to overcoming the absence of socioeconomic data in US medical records. Among 14 240 Black and White men and women who took the KPMCP MHC examination, individual, tract, and block-group measures of social class and education provided highly comparable estimates of association with four diverse health characteristics known to exhibit marked social class and race/ethnic gradients[2,7,25–27]: hypertension, height, smoking, and number of full-term pregnancies.

Beyond this, the use of census data permitted assessing the effect of respondent bias introduced by missing individual-level data, and also allowed contextual analyses to be conducted. These analyses indicated that census data may augment, and not simply approximate, individual-level socioeconomic data. Lastly, the census-based methodology provided a straightforward and relatively inexpensive way of ascertaining whether the KPMCP membership is representative of the surrounding general population, with the findings recently corroborated by the 1989 to 1991 California Risk Factor Surveillance Study (N. Gordon, written communication, February 1991).

Comparison with Other Studies Using Census-Derived Data

To date, very few US studies have compared individual-level and census-derived socioeconomic indicators as correlates of health status.[1,8,21] The greater heterogeneity among residents of census tracts as compared with block groups, however, has long been of concern,[38,39] with some evidence suggesting that block-group data can identify pockets of poverty or affluence not apparent at the tract level.[40] Also bolstering this study's findings are the results of a related pilot project carried out in Alameda County, California.[21] This study of 51 Black and 50 White women found that census block-group measures of social class and poverty closely approximated individual-level measures as correlates of women's reproductive histories, whereas comparable data from the tract level performed less well. Contextual analyses likewise indicated the importance of categorizing women by both individual-level and block-group–level socioeconomic characteristics.[21]

Additional evidence regarding contextual effects and health status stems from another Alameda County study that found that persons who lived in poverty areas were 1.7 times more likely to die during 9 years of follow-up than were persons living in nonpoverty areas, even after controlling for a large number of demographic, physiological, and psychosocial variables.[41] Other US research has observed contextual effects for such diverse phenomena as voting behavior,[42–43] students' test scores,[44] adult participation in community-based organizations,[45] and social problems associated with residential segregation[46] (e.g., Black professionals in the United States are more likely to live in working-class neighborhoods than their White peers, such that analyses "controlling" for only individual-level class differences fail to capture important neighborhood-related differences in socioeconomic position).

TABLE 3—Comparison of Complete and Total Models

Dependent Variable	Model characteristics: Type of Regression Model	Models Adjusted for	Block-Group Model Level	Number in Model: Total	White	Black	Risk Estimates[a] (95% Confidence Intervals): Race	Race + Class + Education
HBP[b]	Logistic	Age Gender Quetelet index	Complete	8216	4960	3256	1.9 (1.7, 2.1)	1.7 (1.5, 1.9)
			Total	14 159	8623	5536	1.7 (1.5, 1.8)	1.5 (1.4, 1.7)
			Increase, %[c]	72.3	73.9	70.0		
Height, cm	Linear	Age Gender	Complete	8241	4975	3266	−0.4 (−0.7, −0.1)	0.0 (−0.3, 0.4)
			Total	14 198	8645	5553	−0.4 (−0.6, −0.2)	0.0 (−0.2, 0.2)
			Increase, %[c]	72.3	73.8	70.0		
Smokes	Logistic	Age Gender	Complete	4247	2554	1703	2.1 (1.8, 2.4)	1.8 (1.6, 2.1)
			Total	5365	3228	2137	2.1 (1.8, 2.3)	1.8 (1.6, 2.0)
			Increase, %[c]	26.3	26.4	25.5		
NFTP	Linear	Age	Complete	4467	2535	1932	0.7 (0.6, 0.8)	0.5 (0.4, 0.7)
			Total	6035	3514	2521	0.6 (0.5, 0.7)	0.4 (0.3, 0.6)
			Increase, %[c]	35.1	38.6	30.5	30.5	

Dependent Variable	Risk Estimates[a] (95% Confidence Intervals): Race + Class + Education
HBP[b]	1.1 (0.9, 1.3)
	1.0 (0.9, 1.2)
Height, cm	−0.7 (−1.0, −0.3)
	−0.6 (−0.9, −0.3)
Smokes	1.2 (1.0, 1.4)
	1.2 (1.1, 1.4)
NFTP	0.2 (0.0, 0.3)
	0.2 (0.0, 0.3)

Dependent Variable	Model characteristics: Type of Regression Model	Models Adjusted for	Block-Group Model Level	Number in Model: Total	White	Black	Risk Estimates[a] (95% Confidence Intervals): Race	Race + Class + Education
HBP[b]	Logistic	Age Gender Quetelet index	Complete	8216	4960	3256	1.9 (1.7, 2.1)	1.7 (1.5, 1.9)
			Total	14 159	8623	5536	1.7 (1.5, 1.8)	1.5 (1.4, 1.7)
			Increase, %[c]	72.3	73.9	70.0		
Height, cm	Linear	Age Gender	Complete	8241	4975	3266	−0.4 (−0.7, −0.1)	0.0 (−0.3, 0.4)
			Total	14 198	8645	5553	−0.4 (−0.6, −0.2)	0.0 (−0.2, 0.2)
			Increase, %[c]	72.3	73.8	70.0		
Smokes	Logistic	Age Gender	Complete	4247	2554	1703	2.1 (1.8, 2.4)	1.8 (1.6, 2.1)
			Total	5365	3228	2137	2.1 (1.8, 2.3)	1.8 (1.6, 2.0)
			Increase, %[c]	26.3	26.4	25.5		
NFTP	Linear	Age	Complete	4467	2535	1932	0.7 (0.6, 0.8)	0.5 (0.4, 0.7)
			Total	6035	3514	2521	0.6 (0.5, 0.7)	0.4 (0.3, 0.6)
			Increase, %[c]	35.1	38.6	30.5	30.5	

Dependent Variable	Risk Estimates[a] (95% Confidence Intervals): Race + Class + Education
HBP[b]	1.2 (1.0, 1.4)
	1.2 (1.0, 1.4)
Height, cm	−0.4 (−0.8, 0.0)
	−0.5 (−0.8, −0.2)
Smokes	1.2 (1.0, 1.4)
	1.2 (1.0, 1.4)
NFTP	0.2 (0.0, 0.3)
	0.2 (0.0, 0.3)

Note. HBP = high blood pressure; NFTP = number of full-term pregnancies.
[a]Comparison groups:
Race: Black vs White
Class: 66 + % working class vs <66% working class Black group
Education: 25 + % < high school vs. <25% < high school Black group
For linear regression: estimate of unit change in outcome associated with unit change in race, class, and education measures
For logistic regression: estimate of relative risk of outcome associated with unit change in race, class, and education measures
[b]HBP defined as diastolic (≥90 mm Hg) and/or systolic (≥ 145 mm Hg).
[c]Percentage increase in number of persons in total as compared with complete model.

TABLE 4—Association of Race and Contextual Social Class Measures with Dependent Variables

Dependent Variable	Model Characteristics: Type of Regression Model	Number in Model: White	Black	Model Adjusted for	Risk Estimates[a] (95% Confidence Intervals): Race	NWC(I) + <66% WC(BG) vs NWC(I) + 66+% WC (BG)	NWC(I) + <66% WC(BG) vs WC(I) + <66% WC (BG)	NWC(I) + <66% WC(BG) vs WC(I) + 66+% WC (BG)
HBP[b]	Logistic	4960	3256	Age Gender Quetelet index	1.7 (1.5, 1.9)	1.2 (1.0, 1.4)	1.1 (1.0, 1.3)	1.3 (1.1, 1.5)
Height, cm	Linear	4975	3266	Age Gender	0.0 (−0.3, 0.4)	−0.9 (−1.3, −0.4)	−1.0 (−1.4, −0.6)	−1.6 (−2.0, −1.2)
Smokes	Logistic	2544	1703	Age Gender	1.8 (1.6, 2.1)	1.4 (1.1, 1.7)	1.6 (1.3, 2.0)	1.8 (1.5, 2.2)
NFTP	Linear	2535	1932	Age	0.6 (0.4, 0.7)	0.2 (0.0, 0.3)	0.1 (−0.1, 0.3)	0.4 (0.2, 0.6)

Note. NWC(I) = non-working class (individual); <66% WC(BG) = less than 66% working class (block group); 66+% WC(BG) = 66% or more working class (block group); WC(I) = working class (individual); HBP = high blood pressure;NFTP = number of full-term pregnancies.
[a]Comparison groups:
Race: Black vs White
For linear regression: estimate of unit change in outcome associated with unit change in race and contextual class measures
For logistic regression: estimate of relative risk of outcome associated with unit change in race and contextual class measures
[b]HBP defined as diastolic (≥90 mm Hg) and/or systolic (≥ 145 mm Hg).

TABLE 5—Demographic Comparison of Block-Group Race/Ethnic Composition of 1985 Kaiser Permanente (KP) Membership and 1980 General Population (GP): 22 Selected Counties in Northern California

Block-Group Composition	Percentage Living in Block Groups with Specified Race/Ethnic Composition														
	White			Black			Native American			Asian/Pacific Islander			Hispanic Origin[a]		
	KP	GP	KP/GP	KP	GP	KP/GP	KP	GP	KP/GP	KP	GP	KP/GP	KP	GP	KP/GP
<20%	4.1	3.7	1.11	88.5	90.7	0.98	100.0	100.0	1.00	89.7	91.4	0.98	84.1	81.4	1.03
20%–39%	4.8	4.4	1.09	4.5	3.7	1.22	0.0	0.0	...	6.1	6.6	1.23	11.5	12.4	0.93
40%–59%	10.0	9.2	1.09	2.4	2.0	1.20	0.0	0.0	...	1.9	1.6	1.19	3.4	4.4	0.77
60%–79%	23.1	21.4	1.08	2.1	1.7	1.24	0.0	0.0	...	0.2	0.2	1.00	0.9	1.5	0.60
80+%	58.0	61.3	0.95	2.4	1.9	1.26	0.0	0.0	...	0.1	0.2	0.50	0.1	0.3	0.33
Total	100.0	100.0		99.9	100.0		100.0	100.0		100.0	100.0		100.0	100.0	

Note. The sample sizes were 1.6 million for KP and 7.4 million for GP. The 22 counties were Alameda, Contra Costa, El Dorado, Fresno, King, Madera, Marin, Napa, Placer, Sacramento, San Francisco, San Joaquin, San Mateo, Santa Clara, Santa Cruz, Solano, Sonoma, Stanislaus, Sutter, Tulare, Yolo, and Yuba.
[a]Can be of any race.

TABLE 6—Demographic Comparison of Block-Group Socioeconomic Characteristics of 1985 Kaiser Permanente (KP) Membership and 1980 General Population (GP): 22 Selected Counties in Northern California[a]

Demographic Variable	Block-Group Composition	Percentage Living in Block Groups with Specified Socioeconomic Composition		KP/GP Ratio
		KP	GP	
Class	<86% working class	39.7	42.4	0.94
	≥86% working class	60.3	57.6	1.05
Poverty	<20% below poverty	91.3	86.8	1.05
	≥20% below poverty	8.7	13.2	0.66
Education	<25% adults less than high school	66.2	61.0	1.09
	≥25% adults less than high school	33.8	39.0	0.87

Note. The sample sizes were 1.6 million for KP and 7.4 million for GP. The 22 counties were Alameda, Contra Costa, El Dorado, Fresno, King, Madera, Marin, Napa, Placer, Sacramento, San Francisco, San Joaquin, San Mateo, Santa Clara, Santa Cruz, Solano, Sonoma, Stanislaus, Sutter, Tulare, Yolo, and Yuba.

Several British investigations likewise have found that area-based measures of deprivation and individual-level social class data detect comparable social gradients in health outcomes.[47–53] As noted by Carstairs and Morris,[48] and by Alexander et al.,[50] these area-based measures possess the additional advantage of being able to evaluate the relationship between health status and socioeconomic position among persons not easily classified by traditional occupation-based class measures (e.g., children, students, and adults not in the active labor force). In these studies, area-based measures of deprivation also revealed intraclass mortality gradients,[49–51] providing additional evidence of contextual effects.

Potential Biases Affecting the Census-Based Methodology

The census-based methodology is not without flaws. First, it requires that individuals have a residential address, and that this address be located within a census-defined tract or block group. In this investigation, 17% of the KPMCP membership could not be geocoded to the block-group level. Despite their similarity in age and gender composition, members with and without block-group data may have differed with regard to other sociodemographic characteristics (e.g., urban vs rural residence), thereby introducing bias. Manual searches on addresses not geocoded by computerized procedures, however, can reduce the percentage of persons not assigned a block-group number to under 5%.[14,15,21] Moreover, beginning with the 1990 census, block-group codes were assigned to the entire nation.[54]

Second, census data might not accurately characterize the demographic context of study subjects because of the decennial nature of census data[55] and also because of the undercount, which chiefly affects poor people and people of color.[55–58] Because population growth and migration can alter a neighborhood's composition,[55] this census-based methodology should be used only for residential addresses falling within 5 years of the closest census. The undercount, in turn, most likely would dilute the association between health status and block-group socioeconomic measures, since it produces conservative estimates of the number and hence proportion of poor persons and people of color.[55–58]

A third concern pertains to ecologic fallacy, which can produce inflated estimates of individual-level associations.[22–23] Although this study used grouped data to characterize individuals, the census-based methodoiogy presented here is not affected by the "classic" type of ecologic fallacy, in which both the dependent and independent variables are grouped data and underlying factors associated with the grouping process confound the results.[22–23] Instead, individual-level dependent variables were analyzed in conjunction with census-based independent variables. To minimize heterogeneity in the census unit, however, it may be preferable to use block-group rather than tract data, if available.[21,38–40]

A final caveat applies to the contextual analyses. Only a limited number of socioeconomic and health-related variables were used in these analyses, since the purpose was to evaluate the association between different socioeconomic factors and health characteristics, as opposed to elucidating the pathways through which these factors exert their effects. Yet, as noted by some critics, if individual-level models are not fully specified, seemingly significant contex-

tual effects may be spurious, with the contextual variables only "explaining" variance that could be better accounted for by additional individual-level data.[37,59] Even so, the plausible nature of the observed contextual trends suggests that this technique may provide a useful means by which to avoid "individualistic fallacy,"[22] that is, the assumption that individual-level data are sufficient to explain social phenomena, including a population's level of health and disease.[60]

Significance and Usefulness of the Census-Based Methodology

The importance of validating this census-based approach to measuring socioeconomic position is underscored by the numerous US studies that, in the absence of individual-level social class data, have used census-derived data from people's immediate neighborhoods in conjunction with individual-level health data to describe, analyze, or control for social gradients in various health outcomes.[9–20] These include investigations regarding race/ethnic differences in cancer incidence and survival,[9–16] homicide,[17] and childhood diseases,[18] as well as studies examining intraurban variation in mortality.[19,20] All have observed significant associations between people's health status and the socioeconomic conditions of the neighborhoods in which they live, and all have expressed concerns regarding the use of census-derived data. The results of this study and comparable research[21,47–53] indicate that these prior findings most likely are legitimate and probably underestimate the effect that would have been observed were individual-level social class data available.

In sum, the census-based methodology presented in this study provides a valid and useful approach to overcoming the absence of socioeconomic data in most US medical records. Because census data are readily available to all US public health researchers and geocoding is not expensive, this methodology easily can be used with any existing data set that contains residential addresses. By offering a measure of socioeconomic position that is applicable to all persons, regardless of age, gender, or employment status, this census-based approach also can greatly assist in both reducing and evaluating response bias resulting from incomplete or inaccurate individual-level socioeconomic information. Other possible uses include ascertaining whether a study sample is representative of the general public, whether a given program is reaching its intended population, and whether socioeconomic factors contribute to small-area variation in health status and health service use patterns. Lastly, in those cases where individual-level data are available, this methodology permits use of contextual analysis, thereby offering new avenues to investigate the diverse routes by which social gradients in disease are produced and maintained.

Apart from raising new questions about social context, the most important contribution of the census-based methodology may be in furnishing US researchers with the means to construct age-specific incidence, prevalence, and mortality rates, stratified by consistent area-based measures of social class. This can be accomplished for most disease outcomes, since denominator data typically are census derived and can be classified according to the same block-group measures.[14] By focusing attention on social gradients in disease within, and not only between, different race/ethnic groups within the United States, such data could potentially reinvigorate, if not reorient, efforts to understand as well as eliminate socially determined disparities in health. □

Acknowledgments

This project was funded by National Cancer Institute contract N01-CP-95606.

This paper was presented at the 118th Annual Meeting of the American Public Health Association, October 1990, New York City.

Thanks to Frank Many for help in obtaining the 1980 census data, to Donna Wells for programming assistance, to Bruce Fireman for suggestions regarding the statistical analyses, and to Gary Friedman and Robert A. Hiatt for their helpful comments and criticisms.

References

1. Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev.* 1988;10:87–121.
2. US Department of Health and Human Services. *Health Status of Minorities and Low Income Groups.* Washington, DC: US Government Printing Office; 1985. DHHS publication no. (HRSA) HRS-P-DV 85-1.
3. Sydenstricker E. *Health and Environment.* New York, NY: McGraw-Hill; 1933.
4. Susser M, Watson W, Hopper K. *Sociology in Medicine.* 3rd ed. New York, NY: Oxford University Press; 1985.
5. Black D, Morris JN, Smith C, Townsend P. *Inequalities in Health: The Black Report.* Harmondsworth, England: Penguin Books; 1985.
6. Smith GD, Bartley M, Blane D. The Black report on socioeconomic inequalities in health 10 years on. *Br Med J.* 1990;301:373–377.
7. Marmot MG. Social inequalities in mortality: the social environment. In: Wilkinson RG, ed. *Class and Health: Research and Longitudinal Data.* London, England: Tavistock; 1986:21–33.
8. Morgenstern H. Socioeconomic factors: concepts, measurements, and health effects. In: Ostfeld AM, Eaker ED, eds. *Measuring Psycho-Social Variables in Epidemiologic Studies of Cardiovascular Disease.* Bethesda, Md: National Institutes of Health; 1985:3–35. NIH publication no. 85-2270.
9. Devesa SS, Diamond EL. Association of breast cancer and cervical cancer incidence with income and education among Whites and Blacks. *JNCI.* 1980;65:515–528.
10. Devesa SS, Diamond EL. Socioeconomic and racial differences in lung cancer incidence. *Am J Epidemiol.* 1983;118:818–831.
11. Ernster VL, Selvin S, Sacks ST, et al. Prostatic cancer: mortality and incidence rates by race and social class. *Am J Epidemiol.* 1978;107:311–320.
12. Dayal HH, Power RN, Chiu C. Race and socioeconomic status in survival from breast cancer. *J Chronic Dis.* 1982;35:675–683.
13. White E, Daling JR, Norsted TL, et al. Rising incidence of breast cancer among young women in Washington State. *JNCI.* 1987;79:239–243.
14. Krieger N. Social class and the Black/White crossover in the age-specific incidence of breast cancer: a study linking census-derived data to population-based registry records. *Am J Epidemiol.* 1990;131:804–814.
15. Bassett MT, Krieger N. Social class and Black-White differences in breast cancer survival. *Am J Public Health.* 1986;76:1400–1403.
16. Savage D, Lindenbaum J, Ryzin JV, et al. Race, poverty, and multiple myeloma. *Cancer.* 1984;54:3085–3094.
17. Centerwall BS. Race, socioeconomic status, and domestic homicide, Atlanta, 1971–72. *Am J Public Health.* 1984;74:813–815.
18. Wise PH, Kotelchuck M, Wilson ML, Mills M. Racial and socioeconomic disparities in childhood mortality in Boston. *N Engl J Med.* 1985;316:360–366.
19. Jenkins CD, Tuthill RW, Tannenbaum SI, Kirby CI. Zones of excess mortality in Massachusetts. *N Engl J Med.* 1977;296:1354–1356.
20. Yeracaris CA, Kim JH. Socioeconomic differences in selected causes of death. *Am J Public Health.* 1978;68:342–351.
21. Krieger N. Women and social class: a methodological study comparing individual, household, and census measures as predictors of Black/White differences in reproductive history. *J Epidemiol Community Health.* 1991;45:35–42.
22. Alker HR Jr. A typology of ecological fallacies. In: Doggan M, Rokkan S, eds. *Social Ecology.* Cambridge, Mass: MIT Press; 1969:69–86.
23. Selvin S. Two issues concerning the analysis of grouped data. *Eur J Epidemiol.* 1987;3:284–287.
24. Kaplan CP, Van Valey TL. *Census '80: Continuing the Fact Finding Tradition.* Washington, DC: US Government Printing Office; 1980.
25. Syme SL, Oakes TW, Friedman GD, Feldman R, Siegelaub AB, Collen M. Social

class and racial differences in blood pressure. *Am J Public Health.* 1974;64:619–620.
26. Walker M, Shaper AG, Wannamethee G. Height and social class in middle-aged British men. *J Epidemiol Community Health.* 1988;42:299–303.
27. National Survey of Family Growth. *Socioeconomic Differentials and Trends in the Timing of Births.* Hyattsville, MD: National Center for Health Statistics; 1981. DHHS publication no. (PHS) 81-1982. (Vital and health statistics; series 23; no. 6).
28. Wright EO, Costello C, Hacken D, Sprague J. The American class structure. *Am Sociol Rev.* 1982;47:709–726.
29. US Bureau of the Census. *Poverty Areas in Large Cities. 1980 Census of the Population, Volume 2, Subject Reports.* Washington, DC: US Government Printing Office; 1985. PC80-2-8D.
30. Congressional Budget Office. *Trends in Family Income: 1970–1986.* Washington, DC: US Government Printing Office; 1988.
31. Beeghley L. Illusion and reality in the measurement of poverty. *Soc Problems.* 1984; 31:322–333.
32. Friedman GD, Selby JV, Quesenberry CP Jr, Armstrong MA, Klatsky AL. Precursors of essential hypertension: body weight, alcohol and salt use, and parental history of hypertension. *Prev Med.* 1988;17:387–402.
33. SAS Institute Inc. *SAS Language Guide for Personal Computers.* Release 6.03 edition. Cary, NC: SAS Institute Inc; 1988.
34. Kleinbaum DE, Kupper LL. *Applied Regression Analysis and Other Multivariable Methods.* Boston, Mass: Duxbury Press; 1978.
35. Breslow NE, Day NE. *Statistical Methods in Cancer Research. Volume I—The Analysis of Case-Control Studies.* Lyon, France: International Agency for Research on Cancer; 1980:191–246.
36. Boyd LH Jr, Iversen GR. *Contextual Analysis: Concepts and Statistical Techniques.* Belmont, Calif: Wadsworth Publishing Co; 1979.
37. Blalock HM Jr. Contextual-effects models: theoretic and methodological issues. *Annu Rev Sociol.* 1984;10:353–372.
38. Foley DL. Census tracts and urban research. *J Am Stat Assoc.* 1953;48:733–742.
39. Myers JK. Note on the homogeneity of census tracts: a methodological problem in urban ecological research. *Soc Forces.* 1954;32:364–366.
40. US Bureau of the Census. *Census Use Study: Health Information System—II.* Report no. 12. Washington, DC: US Government Printing Office; 1976.
41. Haan M, Kaplan G, Camacho T. Poverty and health: prospective evidence from the Alameda county study. *Am J Epidemiol.* 1987;125:989–998.
42. Przeworski A. Contextual models of political behavior. *Political Methodology.* 1974;1:27–61.
43. Segal DR, Meyer MW. The social context of political partisanship. In: Doggan M, Rokkan S, eds. *Social Ecology.* Cambridge, Mass: MIT Press; 1969:217–232.
44. Langbein LI. Schools or students: aggregation problems in the study of achievement. *Eval Stud Annu Rev.* 1977;2:270–298.
45. Bell W. Urban neighborhoods and individual behavior. In: Sherif M, Sherif CW, eds. *Problems of Youth: Transition to Adulthood in a Changing World.* Chicago, Ill: Aldine; 1965:235–264.
46. Erbe BM. Race and socioeconomic segregation. *Am Sociol Rev.* 1975;40:801–812.
47. Curtis SE. Use of survey data and small area statistics to assess the link between individual morbidity and neighborhood deprivation. *J Epidemiol Community Health.* 1990;49:62–68.
48. Carstairs V, Morris R. Deprivation: explaining differences in mortality between Scotland and England and Wales. *Br Med J.* 1989;299:886–889.
49. Carstairs V, Morris R. Deprivation and mortality: an alternative to social class? *Community Med.* 1989;11:210–219.
50. Alexander FE, O'Brien F, Hepburn W, Miller M. Association between mortality among women and socioeconomic factors in general practices in Edinburgh: an application of small area statistics. *Br Med J.* 1987;295:754–756.
51. Morgan M, Chinn S. ACORN group, social class, and child health. *J Epidemiol Community Health.* 1983;37:196–203.
52. Morgan M. Measuring social inequality: occupational classifications and their alternatives. *Community Med.* 1983;5:116–124.
53. Arber S. Social class, nonemployment, and chronic illness: continuing the inequalities in health debate. *Br Med J.* 1987;294:1069–1073.
54. US Bureau of the Census. *Census '90 Basics.* Washington, DC: US Government Printing Office; 1990.
55. White MJ. *American Neighborhoods and Residential Differentiation.* New York, NY: Russell Sage Foundation; 1987.
56. Heer DM, ed. *Social Statistics and the City.* Cambridge, Mass: Joint Center for Urban Studies at the Massachusetts Institute of Technology and Harvard University; 1968.
57. Parsons CW, ed. *America's Uncounted People: Report of the Advisory Committee on Problems of Census Enumeration, Division of Behavioral Sciences, National Research Council.* Washington, DC: US Government Printing Office; 1972.
58. Citro F, Cohen ML, eds. *The Bicentennial Census: New Directions for Methodology in 1990.* Washington, DC: National Academy of Sciences; 1985.
59. Hauser RM. Context and consex: a cautionary tale. *Am J Sociol.* 1970;75:645–664.
60. Rose G. Sick individuals and sick populations. *Int J Epidemiol.* 1985;14:32–38.

Ownership and Performance of Outpatient Substance Abuse Treatment Centers

John R. C. Wheeler, PhD, Huda Fadel, MPH, MAE, and Thomas A. D'Aunno, PhD

ABSTRACT

Background. Little is known about the organization and performance of outpatient substance abuse treatment (OSAT) centers. We examine several performance measures of OSAT units, including clients treated, services provided, revenue sources, financial performance, and access to care, in relation to ownership of the center.

Methods. Data were drawn from a national random sample of 575 OSAT centers (85.8% response rate) participating in a telephone survey conducted in 1988. Analysis of variance by ownership was conducted on each performance measure, with differences subjected to tests of statistical significance.

Results. Descriptive results show that major funding sources differ by ownership. Private for-profit centers generate higher profits, charge higher prices, and achieve higher levels of financial performance than public and not-for-profit centers. Public centers provide better access to care for persons who are unable to pay.

Conclusions. There appear to be substantial and interrelated differences by ownership type in the financing and operation of OSAT units. (*Am J Public Health.* 1992; 82:711–718)

Introduction

Substance abuse is a major health hazard in the United States, accounting for over $2.4 billion in treatment and prevention expenditures in 1989.[1] The National Drug Control Strategy appropriated approximately $925 million for treatment interventions in fiscal year 1990, representing a significant increase in public commitment to reduce drug demand.[2] But despite this high level of investment, little is known about the outpatient substance abuse treatment (OSAT) organizations that provide most of the treatment. Specifically, there is a lack of up-to-date information regarding the nature of care provided in OSAT units, the revenue sources and financial performance of the programs, or even their accessibility to persons in need of services. This paper examines these issues using a nationally representative sample of OSAT organizations.

A brief overview is useful for understanding the substance abuse treatment system. This paper focuses on outpatient providers who currently treat the vast majority (more than 85%) of clients in the substance abuse treatment system.[3] Two treatment modalities predominate: outpatient methadone maintenance and outpatient drug-free therapy. In the first, treatment begins with the use of a synthetic narcotic (methadone) being substituted for the abuse of opioid drugs, enabling the client to taper off gradually and eventually eliminate drug use. In the second, a combination of individual and group counseling, psychotherapy, and supervision is used to eliminate drug use. The provision of both modalities has grown rapidly since the 1960s owing to government and private funding, but drug-free therapy is currently the prevalent treatment mode.

Much of the substance abuse treatment research has focused on clients and particularly on what happens to individuals who undergo treatment. Outpatient substance abuse treatment is less well understood, however, on an organizational level. Although most programs use similar treatment methods (i.e., individual, group, and family therapy) to produce OSAT care, other organizational differences between providers can be expected to result in different modes of operation and performance.[4] The question is, what are some of these organizational characteristics and how do they relate to observed performance in OSAT units?

One potentially important characteristic is the ownership status of the treatment unit—i.e., whether it is owned and operated on a private for-profit (FP), private not-for-profit (NFP), or public basis. The salience of ownership status derives from how that status is connected with the organization's objectives, such as profit maximization, access to care, prestige, patient outcomes, etc. At one theoretical extreme, FP firms operate to maximize the wealth of the owners. At the other extreme, public firms seek to promote social welfare. In between these two extremes, NFP firms seek to achieve both social and private objectives within a private context.[5]

The authors are with the University of Michigan's Institute for Social Research and School of Public Health in Ann Arbor.

Requests for reprints should be sent to John R. C. Wheeler, School of Public Health, University of Michigan, Ann Arbor, MI 48109-2029.

This paper was submitted to the Journal September 18, 1990, and accepted with revisions September 20, 1991.

Economic theory predicts that FP firms will behave differently from NFP or public firms because of the presence of clearly defined owners. These owners and their managers are rewarded for operating efficiently; i.e., excess profits are distributed. Incentives for efficiency tend to be weaker in NFP and public firms because of a lack of clear ownership and the nondistribution of profits.[6] In practice, however, organizational performance may or may not differ, depending not only on organizational objectives but also on incentives provided by funding sources, differences in services provided, characteristics of clients treated, influences of program staff, and other market characteristics.[7]

Many research studies have examined the association between ownership and organizational performance in health care. Several hospital studies have suggested important and substantive differences in the behavior of NFP and FP firms. Findings include the production of costlier and higher priced care in FP hospitals,[8] higher prices and profits in FP hospitals,[9] better access in NFP hospitals,[8,10] and more charity care in NFP hospitals.[8,11] Other studies conclude that NFP hospitals are less efficient than FP hospitals, after adjustment for tax subsidies.[12] Finally, there are still others that report no important differences between the two ownership types.[13] Similarly mixed results have been reported for other types of health care organizations.

In the substance abuse field, the public sector has traditionally assumed major responsibility for the operation of treatment programs. The role of the private sector has been increasing, however—especially that of FP firms. To a great extent, the growth in private sector ownership of OSAT organizations is simply a part of the general trend in health care toward increased private delivery of services.[14,15] It also reflects specific public policies in the 1980s to increase private sector participation in the war on drugs, such as the 1984 national strategy for drug abuse policy.[16] As in other health sectors, however, the relationship between increasing privatization and the effectiveness and efficiency of substance abuse treatment is not well understood. Further study is needed to assess the benefits and potential costs of private sector ownership in this field.

This paper describes OSAT clients and service characteristics, funding for OSAT programs, the programs' financial performance, and the accessibility of their services. Differences in organizational performance that are related to ownership status are highlighted and discussed.

Methods and Measures

Sample

The data for this paper were collected from a national study of OSAT organizations that was supported by the National Institute on Drug Abuse. The study was conducted in the fall of 1988 by the Institute for Social Research at the University of Michigan. A random sample of 670 OSAT units was selected from a total population of 6851 such units in the United States in 1988. The sample was stratified according to treatment modality (methadone or drug free), treatment setting (mental health center, hospital-based program, freestanding facility) and ownership status (private FP, private NFP, and public). Five hundred and seventy-five units participated in the survey, yielding a response rate of 85.8%. The response rate for private units (both FP and NFP) was 88.0%; for public units, 81.7%. Each unit's administrative director and clinical supervisor were asked to complete telephone surveys. Directors provided information about their unit's financial status and services, and clinical supervisors provided information about personnel, client characteristics, and substance abuse problems. All information was collected for the unit's most recent complete fiscal year.

Measures

Organizational performance was represented by four sets of multiple measures (see Appendix).

Analysis Method

Analysis of variance tests were conducted on each performance measure, using ownership status as the stratification variable. Differences in performance across ownership types were then subjected to statistical significance tests.

Results

Clients and Services

Measures that describe and differentiate the units in terms of client and service characteristics are displayed in Table 1. More than 69% of the units are devoted exclusively to drug-free treatment, leaving 31% that provide any amount of methadone therapy. By ownership strata, 74% of NFP, 65% of FP, and 63% of public units are drug-free treatment centers.

Classification of clients in treatment according to their specific substance abuse problem is difficult largely because of prevailing patterns of multiple substance use. Nevertheless, clients can be generally categorized as to whether they are primarily alcohol or drug abusers. The primary problem for more than 40% of clients at treatment units is alcohol abuse whereas 51% are primarily drug abusers. (The remaining 9% are clients who could not be assigned to either primary category.) These patterns do not differ significantly by ownership.

Abuse of specific types of substances, singly or in multiple combinations, is common, even for those clients whose primary problem is alcohol addiction. Treatment units overall reported that, on average, 23% of their clients have problems with heroin abuse whereas 27% abuse cocaine. For-profit units tend to have the highest percentage of clients with heroin problems (34%), followed by public units (25%) and NFP units (20%). A similar pattern also holds for cocaine abuse: the percentage of cocaine abusers is highest in FP units (30%), followed by public units (27%) and NFP units (26%). Public units report the highest percentage of multiple drug abusers (55%) whereas FP units report the lowest (44%); these differences are statistically significant. These measures indicate that the major substance abuse problems presented at OSAT units may be similar in nature. The measures are too broad, however, to make conclusions regarding client case-mix severity or service intensity among different units.

Treatment units vary significantly in terms of the mix of therapy types, length of therapy sessions, and group size. On the whole, individual therapy accounts for 78% of all treatment sessions and 72% of all treatment hours provided. Public units produce the highest proportion of individual therapy sessions and hours, followed by private NFP and FP units. The average lengths of therapy sessions show inconsistent patterns that may reflect underlying differences in product mix among the units. Individual therapy sessions are longest in NFP and public units; however, group therapy sessions are longest in FP units. This may reflect the larger treatment group sizes seen at FP units. It may also indicate other characteristic differences, such as type of group therapy provided. In addition, patients in FP units attend group therapy more times per week than patients at other types of units.

TABLE 1—Measures of Services Offered by Treatment Units for All Centers and by Unit Ownership

Service Measure		Total	Private For-Profit (FP)	(1)	Private Not-For-Profit (NFP)	(2)	Public	(3)	(4)
% of units that are drug free	mean	69.19%	64.86%		73.70%	***	62.50%		**
	SD	0.46	0.48		0.44		0.49		
	n	529	37		308		184		
% of clients, primary prob: alcohol	mean	40.71%	41.08%		40.16%		41.54%		
	SD	30.53	32.20		30.30		30.71		
	n	513	37		297		179		
% of clients, primary prob: drugs	mean	51.01%	54.81%		50.05%		51.81%		
	SD	34.70	37.18		34.49		34.67		
	n	515	37		298		180		
% of clients: heroin abuse	mean	22.78%	34.17%	**	20.13%		24.82%		**
	SD	34.94	42.91		32.51		36.60		
	n	508	36		293		179		
% of clients: cocaine abuse	mean	26.87%	29.51%		26.41%		27.09%		
	SD	22.11	25.29		21.36		22.71		
	n	510	37		293		180		
% of clients: multiple drug use	mean	52.61%	43.81%		52.09%		55.28%	**	
	SD	32.20	33.80		31.11		33.68		
	n	502	37		288		177		
IT as % of total sessions provided	mean	77.98%	73.24%		77.45%		80.06%	*	
	SD	0.20	0.27		0.19		0.19		
	n	506	39		301		166		
IT as % of total hours provided	mean	72.41%	67.94%		71.21%	**	75.64%	*	*
	SD	0.22	0.29		0.22		0.21		
	n	509	38		305		166		
No. of minutes per IT session	mean	53.89	49.46	***	54.76		53.25	*	**
	SD	11.70	15.47		11.41		11.08		
	n	518	37		312		169		
No. of minutes per GT session	mean	97.60	105.83		98.68		94.17	*	
	SD	30.87	42.91		31.27		27.12		
	n	491	30		296		165		
No. of minutes per FT session	mean	74.47	75.69		74.80		73.59		
	SD	33.57	42.29		30.72		36.79		
	n	435	29		264		142		
No. of clients per GT session	mean	9.84	10.67		9.61		10.12		
	SD	4.52	3.70		4.72		4.27		
	n	490	30		295		165		
No. of clients per FT session	mean	2.98	2.93		3.03		2.91		
	SD	1.18	1.57		1.26		0.93		
	n	418	27		254		137		
No. of GT sessions per client per week	mean	1.59	2.03		1.68	**	1.33	**	**
	SD	1.67	2.87		1.67		1.25		
	n	521	38		311		172		

Notes. SD = standard deviation, n = number. IT = individual therapy, GT = group therapy, and FT = family therapy.
(1) Significance level for test of differences between means of private FP and private NFP units.
(2) Significance level for test of differences between means of private NFP and public units.
(3) Significance level for test of differences between means of private FP and public units.
(4) Significance level for test of differences between means of private FP, private NFP, and public units.
*significant at .1.
**significant at .05.
***significant at .01.

Revenue Sources

Approximately two-thirds of OSAT revenue comes from public sources (see Table 2). The bulk of this money comes directly from state government sources, with the remainder coming largely from city and county government. Only a small percentage was reported as direct revenue from the federal government. Approximately one-third of total revenue comes from private sources, principally as client out-of-pocket payments and payments from private insurance companies.

Sources of financing differ significantly by ownership category. Private FP units receive 84% of their revenue from private sources. Most of this money (47%) comes from patient out-of-pocket payments, with private insurance being the second largest source (37%). Only 11% of FP units' revenue comes from government sources. In sharp contrast, public treatment units receive 80% of their revenue from public sources and only 19% from private sources. In NFP units, the mix of revenue sources is similar to that for the study sample as a whole, with approximately two-thirds of revenue coming from public sources and one-third coming from private sources.

The way treatment is funded relates to two important issues. First, financing source differences may reflect systematic differences in the types of clients seen by the treatment units, in terms of clients' social and economic resources. Second, each funding source may provide revenue in several ways. The basis upon which units receive revenue—e.g., fee for ser-

TABLE 2—Percent of Revenue by Source for All Centers and by Unit Ownership

Revenue Source		Total	Private For-Profit (FP)	(1)	Private Not-For-Profit (NFP)	(2)	Public	(3)	(4)
Government	mean	66.13%	10.76%	**	65.11%	**	80.12%	**	**
	SD	0.34	0.21		0.33		0.22		
	n	501	35		305		161		
Federal	mean	4.10%	1.59%		4.03%		4.77%		
	SD	0.14	0.06		0.14		0.16		
	n	503	35		306		162		
State	mean	44.07%	5.36%	**	42.12%	**	55.91%	**	**
	SD	0.36	0.16		0.37		0.32		
	n	506	35		306		165		
Local	mean	18.03%	3.81%	*	19.03%		19.18%	*	*
	SD	0.27	0.12		0.29		0.26		
	n	506	35		306		165		
Private	mean	31.57%	84.38%	**	32.30%	**	18.96%	**	**
	SD	0.31	0.25		0.30		0.21		
	n	474	32		291		151		
Insurance	mean	12.74%	37.32%	**	14.24%	**	4.74%	**	**
	SD	0.24	0.35		0.26		0.11		
	n	494	33		303		158		
Out-of-pocket	mean	17.81%	47.30%	**	16.49%		13.38%	**	**
	SD	0.22	0.36		0.19		0.17		
	n	489	34		300		155		
Donations	mean	1.72%	1.38%		2.40%	**	0.55%		**
	SD	0.05	0.07		0.06		0.02		
	n	496	34		299		163		

Notes. SD = standard deviation, n = number.
(1) Significance level for test of differences between means of private FP and private NFP units.
(2) Significance level for test of differences between means of private NFP and public units.
(3) Significance level for test of differences between means of private FP and public units.
(4) Significance level for test of differences between means of private FP, private NFP, and public units.
*significant at .01.
**significant at .001.

vice, capitation, lump-sum grant, etc.—may provide diverse incentives for the organization of treatment.

Financial Performance

Table 4 presents the financial performance of OSAT operations. For the whole sample, total expenses per therapy hour and per therapy session are $62 and $81, respectively. Expenses for salary and wages are $47 per therapy hour and $58 per therapy session. On average, salary and wages comprise approximately three-quarters of total OSAT expenses—a high percentage being devoted to service provision as opposed to overhead and other expenses.

These financial measures differ significantly by ownership type. Public units have higher costs per therapy hour than either FP or NFP units. For-profit units have lower costs per session than both NFP and public units. In the absence of client outcome data, inferences regarding "efficiency" differences among units cannot be made, nor can inferences regarding cost per client, as services and clients may vary across units.

Average OSAT revenue per therapy hour and per therapy session for the whole sample are $66 and $80, respectively. Analysis by ownership types indicates that FP units receive lower amounts of revenue per unit of output than public and NFP units. On the other hand, FPs typically charge much higher prices for individual, group, and family therapy sessions than either NFP or public units. Because therapy sessions often differ in length, a standardized measure was created to compare prices based on a 60-minute therapy hour. The results (not shown) confirm the finding that FP units charge significantly higher prices than NFP and public programs.

It is not uncommon in health care organizations for prices charged to differ from revenues received. A firm's charges are usually higher than the revenues it receives per unit of service because of discounts granted to third-party payers, uncompensated care, etc. There is, however, an apparent contradiction in the fact that the prices charged by the FP units are relatively higher while the revenues per service received are lower, whereas the relation is reversed for public units, which charge lower prices but receive higher revenues per service. A principal explanation has to do with the way in which units receive revenue. For example, FP units receive most of their revenue from client fees, whereas the public units receive substantial revenue in the form of block grant subsidies. This suggests that pricing behavior may be related to incentives embedded in the way a unit receives its revenue. That is, revenues that are not received from clients or based on organizational performance, such as block grants, allow organizations to charge lower prices to clients.

Profit margin, indicating the business efficiency of unit operations, is the final measure of financial performance (see Table 3). For the sample as a whole, treatment units earn an average profit margin of 2%. However, profit margins differ widely across ownership types. For-profit organizations earn a return of 8% of revenues whereas public units break even. Not-for-profit organizations earn profits of 2% of revenues. Comparison of FP performance with that of other units shows significant differences, whereas compari-

TABLE 3—Measures of Treatment Unit Financial Performance for all Centers and by Unit Ownership

Efficiency Measure		Total	Private For-Profit (FP)	(1)	Private Not-For-Profit (NFP)	(2)	Public	(3)	(4)
Total expense per therapy hour	mean	$ 61.80	$58.28		$ 57.12	***	$ 71.08		**
	SD	52.42	67.75		47.47		56.84		
	n	451	28		274		149		
Total expense per therapy session	mean	$ 80.69	$66.44		$ 80.31		$ 84.20		
	SD	73.74	74.12		76.30		68.93		
	n	464	31		278		155		
Salary and wages expense per therapy hour	mean	$ 46.74	$35.90		$ 44.25	*	$ 53.43	*	
	SD	51.02	35.18		48.95		56.56		
	n	453	29		275		149		
Salary and wages expense per therapy session	mean	$ 58.30	$40.91		$ 57.06		$ 64.10	*	
	SD	61.64	39.71		55.96		73.41		
	n	463	32		276		155		
Total revenue per therapy hour	mean	$ 65.56	$57.63		$ 62.46	*	$ 73.12		
	SD	62.78	52.62		61.49		66.71		
	n	473	32		289		152		
Total revenue per therapy session	mean	$ 79.86	$64.87		$ 79.75		$ 83.40		
	SD	71.21	59.00		72.12		71.97		
	n	481	35		289		157		
Price per IT session	mean	$ 33.68	$51.41	****	$ 34.84	***	$ 27.03	****	****
	SD	24.92	27.69		23.82		24.05		
	n	424	32		261		131		
Price per GT session	mean	$ 21.50	$42.00	****	$ 23.18	****	$ 14.02	****	****
	SD	21.10	30.80		20.90		14.79		
	n	412	27		254		131		
Price per FT session	mean	$ 35.11	$54.63	***	$ 36.20	**	$ 28.86	****	****
	SD	30.70	33.69		26.61		35.47		
	n	361	24		223		114		
Profit margin	mean	1.77%	7.78%	*	2.31%		–0.21%	**	**
	SD	0.16	0.25		0.16		0.13		
	n	454	27		270		157		

Notes. SD = standard deviation, n = number, IT = individual therapy, GT = group therapy, and FT = family therapy.
(1) Significance level for test of differences between means of private FP and private NFP units.
(2) Significance level for test of differences between means of private NFP and public units.
(3) Significance level for test of differences between means of private FP and public units.
(4) Significance level for test of differences between means of private FP, private NFP, and public units.
*significant at .1.
**significant at .05.
***significant at .01.
****significant at .001.

sons of profitability between NFP units and public units do not.

Accessibility

Table 4 summarizes the extent to which OSAT units promote access to and equity of service provision. In the full sample, an average of 26% of clients were unable to pay for the care they received; the measure ranges from 8% for the FP units to 32% for the public units. The percentage of clients seen who pay a reduced fee is 47% for all units. Again, FP units treat a much smaller percentage of patients under reduced fee arrangements. Taken together, the percentage of clients who either pay nothing or pay a reduced fee for care received is 82% for public units, 74% for NFP units, and 23% for FP units.

Only 2% of all patients are turned away because they are unable to pay for services. Public units turn away virtually no clients whereas FP units turn away 7% of potential clients.

The length of waiting lists for care, in terms of the numbers of clients in the queue and days of waiting time, is an important measure of whether the existing system is meeting patient needs. The average waiting list for all units is 16 clients and 10 days to begin treatment. Mean waiting lists are somewhat longer for public units and somewhat shorter for FP units although the differences are not significant. An additional analysis (not shown), which relates the number of potential clients on the waiting list to the size or capacity of the unit, does not indicate ownership differences.

Discussion

The OSAT system consists of diverse provider organizations. This paper has categorized treatment units by ownership and found that significant differences characterize the units in terms of clients treated, services offered, funding sources, financial performance, and accessibility.

One of the most critical results is the significant difference in major funding sources by ownership type. This finding is especially important because of the relationship between funding sources and several organizational outcomes—that is, types of clients treated, service mix provided, financial performance, and extent of access to services. For example, even though OSAT programs may not differ significantly in terms of treatment modality or major types of substance problems treated (Table 1), differences in their major revenue sources (Table 2) suggest that the clientele may, in fact, differ between the units. Programs receiving most of their revenue from private insurance or out-of-

TABLE 4—Measures of Access Provided by Treatment Units for All Centers and by Unit Ownership

Access Measure		Total	Private For-Profit (FP)	(1)	Private Not-For-Profit (NFP)	(2)	Public	(3)	(4)
% of clients unable to pay	mean	26.13	8.03	***	25.07	**	32.21	***	***
	SD	28.52	16.94		28.04		29.60		
	n	501	38		298		165		
% of clients paying a reduced fee	mean	46.65	14.54	***	49.24		49.52	***	***
	SD	37.02	20.08		37.27		36.22		
	n	504	39		300		165		
% of clients turned away	mean	1.92	7.44	***	2.10	*	0.43	***	***
	SD	8.96	12.64		10.35		2.85		
	n	509	36		303		170		
No. of clients on waiting list	mean	15.93	7.67		14.37		20.18		
	SD	48.16	23.52		33.45		68.32		
	n	520	36		303		181		
No. of days on waiting list	mean	9.82	7.11		9.28		11.25		
	SD	18.19	22.49		16.12		20.37		
	n	519	36		301		182		

Notes. SD = standard deviation, n = number
(1) Significance level for test of differences between means of private FP and private NFP units.
(2) Significance level for test of differences between means of private NFP and public units.
(3) Significance level for test of differences between means of private FP and public units.
(4) Significance level for test of differences between means of private FP, private NFP, and public units.
*significant at .05.
**significant at .01.
***significant at .001.

pocket payments are likely to be treating a more affluent clientele. Differences in clients' economic and social resources have been associated with differences in drug problem severity.[17] However, detailed information on client severity and outcomes are not available from the current study. Further analysis that controls for patient characteristics needs to be done.

Associated with differences in funding sources are alternative payment methods—i.e., how units receive revenue. Certain funding sources, such as some third-party payers, often impose constraints on the types or amount of services delivered or on the types of personnel used to produce services. In addition, the basis upon which revenue is received—e.g., fee for service, capitation, or lump-sum grant—varies by funding source. Incentives embedded in various payment arrangements have implications for service organization and allocation. For example, it has been found that per visit reimbursement encourages more sessions per course of treatment on average than a lump-sum reimbursement arrangement in OSAT centers.[18]

Additional findings from the study further indicate that ownership is an important determinant of OSAT organizational outcomes. It was reported that FP units are operated at a higher level of profit than are NFP and public units (Table 3). There are several possible explanations for this higher profitability, the most obvious being that FP units operate under a different set of objectives than do public and NFP units. It is also possible that FP units provide services more efficiently; however, this cannot be determined without controlling for client outcomes and product mix. Finally, as mentioned above, the fact that FP units rely heavily on private revenue sources may also be a major influence on their behavior and outcomes.

Because public and NFP units accept a higher percentage of nonpaying and reduced-fee clients (Table 4), they appear to do more to facilitate access to care for persons whose means of payment are limited. Such behaviors are consistent with organizational and economic theories that state that these types of firms exist to provide public or social goods that the private FP sector may underprovide.

There appears to be substantial differences in the financing and operation of OSAT units related to ownership type. Generally speaking, these differences are more striking than those reported in the literature for other health care providers. What is not clear is why these differences exist. In that respect, this paper probably raises more questions than it answers. One obvious set of questions concerns whether the observed differences in revenue sources, finances, accessibility, and service types are directly attributable to ownership or are more closely tied to other characteristics that happen to be associated with ownership. For example, are differences in financial performance due to the fact that FP organizations may have objectives that differ from those of public units or are they due to differences in the constraints faced by the two ownership types relating to primary revenue sources, clients served, or a combination of the above?

Another important set of questions relates to the meaning of other differences found. For instance, does the fact that FP units provide a lower percentage of individual therapy have any relationship to the overall quality of care delivered? Likewise, do the pricing differentials for services between units by ownership type reflect underlying differences in the types of therapy delivered, or do they simply reflect efficiency differences? Future research will be necessary to answer these and similar important questions. □

Acknowledgments

This study was sponsored by the National Institute on Drug Abuse Grant DA03272.

The authors wish to express their gratitude to David M. Klingel, William C. McCaughrin, Richard H. Price, and two anonymous reviewers who offered conceptual and technical assistance in the development of this manuscript as well as thoughtful reviews of earlier drafts.

APPENDIX—Measures of Organizational Performance

Clients and Services

Treatment modality: percent of drug-free units[a]
Primary substance abuse problem[b]
- Percent of clients whose primary problem is alcohol abuse
- Percent of clients whose primary problem is drug abuse

Abuse of other drugs[c]
- Percent of clients abusing heroin
- Percent of clients abusing cocaine
- Percent of clients abusing multiple drugs

Service mix
- Individual therapy as percent of total therapy sessions
- Individual therapy as percent of total therapy hours

Service characteristics
- Minutes per individual therapy session
- Minutes per group therapy session
- Minutes per family therapy session
- Clients per group therapy session
- Number of group therapy sessions per client per week
- Number of clients per family therapy session

Revenue Sources

Government
- Federal
- State
- Local

Private
- Insurance[d]
- Out-of-pocket payments
- Donations

Notes.
1. All revenue sources are expressed as a percentage of total revenue.
2. Federal grants programs administered by state governments were counted as state government revenue.

Financial Performance

Total expenses per therapy hour
Total expenses per therapy session
Salary and wages expenses per therapy hour
Salary and wages expenses per therapy session
Total revenue per therapy hour
Total revenue per therapy session
Price per individual therapy session
Price per group therapy session
Price per family therapy session
Profit margin[e]

Notes.
1. Expenses, revenues, salary, and wages are specific to OSAT care.
2. Therapy hours are total number of hours of individual, group, and family therapy.
3. Therapy sessions are total number of sessions of individual, group, and family therapy.

Accessibility

Percent of clients unable to pay for care received
Percent of clients paying reduced fee for care received
Percent of clients turned away because of inability to pay
Number of potential clients on waiting list for treatment
Number of days on waiting list for treatment.

[a]Programs providing any amount of methadone therapy were classified as methadone units.
[b]Refers to the *primary* substance abuse problem treated. The sum of alcohol and drug abuse is 100%.
[c]Refers to abuse of drugs independent of the primary substance abuse problem. May be a primary or secondary problem. The sum of the percent of clients abusing heroin minus the percent of clients abusing multiple drugs is *not* equal to the percent of clients whose primary problem is drug abuse.
[d]Includes commercial insurance and Blue Cross/Blue Shield plans. Medicare payments are categorized as federal revenue, and Medicaid payments are categorized as state revenue.
[e]Defined as net income divided by total revenue.

References

1. Butynski W, Canova D, Rada J. State resources and services related to alcohol and other drug abuse problems: fiscal year 1989. National Association of State, Alcohol and Drug Abuse Directors, Inc.; Washington, DC: August 1990.
2. The White House. *National Drug Control Strategy.* Washington, DC: US Govt Printing Office; 1989.
3. Office of Technology Assessment. *The Effectiveness of Drug Abuse Treatment: Implications for Controlling AIDS/HIV Infection.* Background paper in series on AIDS-related issues. Washington, DC: Office of Technology Assessment; September 1990.
4. D'Aunno TA, Price RH. Organizational adaptation to changing environments: community mental health and drug abuse services. *Am Behav Scientist.* 1985;28:669–683.
5. Weisbrod BA. *The Voluntary Nonprofit Sector.* Lexington, Mass: Lexington Books; 1977.
6. Hansmann H. Economic theories of nonprofit organization. In Powel, WW, ed. *The Nonprofit Sector: A Research Handbook.* New Haven, Conn: Yale University Press; 1987.
7. Schlesinger M, Marmor TR, Smithey R. Nonprofit and for-profit medical care: shifting roles and implications for health policy. *J Health Polit Policy Law.* 1987;12:427–457.
8. Gray HB, ed. *For-Profit Enterprise in Health Care.* Washington, DC: National Academy Press; 1986.
9. Renn SC, Schramm CJ, Watt JM, Derzon RA. The effects of ownership and system affiliation on the economic performance of hospitals. *Inquiry.* 1985;22:219–236.
10. Lee AJ, Weisbrod BA. Collective goods and the voluntary sector: the case of the hospital industry. In: Weisbrod BA. *The Voluntary Nonprofit Sector.* Lexington, Mass: Lexington Books; 1977.
11. Lewin LS, Eckels TJ, Miller LB. The provision of uncompensated care by not-for-profit hospitals. *N Engl J Med.* 1988; 318:1212–1215.
12. Herzlinger R, Krasker WS. Who profits from nonprofits? *Harvard Business Rev.* 1987;65:93–102.
13. Sloan FA, Vraciu RA. Investor-owned and not-for-profit hospitals: addressing some issues. *Health Aff.* 1983;12:25–37.
14. Starr P. *The Social Transformation of American Medicine.* New York, NY: Basic Books; 1982.
15. Jaffee JH. Evaluating drug abuse treatment: a comment on the state of the art. In: Tims FM, Ludford JP, eds. *Drug Abuse Treatment Evaluation: Strategies, Progress and Prospects.* Research Monograph 51. Rockville, Md: National Institute on Drug Abuse; 1984. US Dept of Health and Human Services publication ADM 84–1329.
16. Drug Abuse Policy Office, White House

Office of Policy Development. *1984 National Strategy for Prevention of Drug Abuse and Drug Trafficking*. Washington, DC: US Govt Printing Office; 1984.
17. McLellan TA. Patient characteristics associated with outcome. In: Cooper JRC, Altman F, Brown BS, Czechowicz D, eds. *Research on the Treatment of Narcotic Addiction: State of the Art*. Treatment Research Monograph Series. Rockville, Md: National Institute on Drug Abuse; 1983. US Dept of Health and Human Services publication ADM 83–1281.
18. Edlund M, Wheeler JRC, D'Aunno TA. Payment systems and payment incentives in outpatient substance abuse treatment. *J Public Budgeting Financial Manage*. 1991. In press.

Israeli/Palestinian International Conference on Water

The first Israeli/Palestinian International Academic Conference on Water will be held October 20 to 23, 1992, in Jerusalem/Bethlehem at the Tantur Conference Center. The Conference program will include lectures, roundtables, and discussion groups on the water resources conflicts between Israelis, Palestinians, and others in the area and on possible solutions.

Specific topics to be discussed: history of the water conflicts of the region; present status of water resources in the region; role of international water law; technical options for increased water resources; regional projects for importing water and/or desalination; nonconventional water sources/water recycling, cloud seeding, fossil water, brackish water; efficient water use and conservation; environmental quality management of shared water resources; joint commissions for monitoring, control, and management of shared international water resources and projects; technical assistance; international financing; and other topics.

The conference language is English, and the conference is open to scientists, engineers, and policymakers from all countries.

Low-cost hotel facilities are available at the Tantur Conference Center, which is conveniently located on the main road between Jerusalem and Bethlehem.

For further information write to either one of the Conference Cochairs and sponsoring organizations: Cochair Dr. Jad Isaac, Applied Research Institute, PO Box 860 Bethlehem, Via Israel (Fax: 972-2-741889); Cochair Professor Hillel Shuval, Truman Institute for the Advancement of Peace, The Hebrew University of Jerusalem, Jerusalem, Israel (Fax: 972-2-322545/666804).

Public Health Briefs

Home Apnea Monitoring and Disruptions in Family Life: A Multidimensional Controlled Study

Elizabeth Ahmann, ScD, RN, Louise Wulff, RN, ScD, and Robert G. Meny, MD

ABSTRACT

We used data from telephone interviews and mailed questionnaires to examine 12 aspects of family life among 93 families with infants considered at high risk for sudden infant death syndrome and on home apnea monitors and a matched comparison group with infants not requiring monitoring. Using logistic regression to control confounding variables, we found that case mothers were at an increased risk of poor health, but we found no other significant differences in family life between the two groups. (*Am J Public Health.* 1992;82:719–722)

Introduction

The presumed association between apnea and sudden infant death syndrome (SIDS) has led to the use of home apnea monitors for diagnostic and preventive purposes. According to the Congressional Office of Technology Assessment, as many as 45 000 infants are on home apnea monitors,[1] a prevalence of 11 to 12.4 monitored infants per 1000 live births.[2]

Despite its prevalence, home apnea monitoring remains controversial, in part because of concern about its possible detrimental effects on the family.[3–5] While descriptive studies[6–10] have suggested effects including parental fatigue, anxiety, social isolation, and depression, analytic studies[11–14] have had varied and conflicting results.

To provide a more comprehensive and well-controlled examination of family life than in previous studies of monitored infants, we conducted a cross-sectional study of family members' health and behavior; intrafamilial roles, relationships, and functioning; and family–social relationships.

Methods

All case families were participants in two monitoring programs providing support and follow-up similar to that recommended by the National Institutes of Health.[5] We collected data from 93 families of infants considered at high risk for SIDS and on home monitors (80.9% response rate) and 86 comparison families with infants not requiring monitoring (84.3% response rate). Comparison families, identified from birth/nursery records, were matched to cases on infant age, race, and sex, as well as maternal marital status, parity (1 vs >1), and health insurance status. Institutional review boards at each site approved the study.

A higher percentage of case respondents (63%) than nonrespondents (31.8%) had commercial insurance ($\chi^2 = 7.379$, $df = 2$, $P = .025$). Respondents and nonrespondents did not differ significantly on infant age, gestational age, race, sex, reason for monitoring, or maternal marital status or parity.

Telephone interviews with mothers provided data on potentially confounding variables (Tables 1 and 2). Questionnaires mailed separately to mothers and fathers included scales measuring 12 dependent family life variables: the Parental Depression, Parental Health, Sense of (parenting) Competence, Attachment (to infant), Restrictions of (parenting) Role, Relationship with Spouse, and Social Isolation subscales of Abidin's Parenting Stress Index[15] (Chronbach's alpha reliabilities[16] ranged from 0.58 to 0.99); the Kansas Marital Satisfaction Scale[17] (alpha of 0.95); the frequency scale of Eyberg's Child Behavior Inventory[18] (sibling behavior) (alpha of 0.99); and three scales developed specifically for this study to measure the quality of both the parent–

Elizabeth Ahmann is a consultant in Washington, DC. At the time of the study Louise Wulff was with the Department of Maternal and Child Health, School of Hygiene and Public Health, The Johns Hopkins University, Baltimore, Md. She is now with the National Fetal-Infant Mortality Review Project, American College of Obstetricians and Gynecologists. Robert G. Meny is with the SIDS Institute, Department of Pediatrics, University of Maryland Medical School, College Park.

Requests for reprints should be sent to Elizabeth Ahmann, ScD, 3417 22nd Street, NE, Washington, DC 20018.

This paper was submitted to the Journal August 21, 1990, and accepted with revisions September 16, 1991.

TABLE 1—Distribution of Case and Comparison Mothers on Potentially Confounding Variables

	Case Mothers (n = 93)		Comparison Mothers (n = 86)	
	No.	%	No.	%
Infant medical problems[a]				
No	56	60.2	62	72.1
Yes	37	39.8	24	27.9
Infant developmental delay				
No	83	90.2	86	100.0
Yes	9	9.8	0	...
Days infant hospitalized[b]				
≤5	39	41.9	62	72.1
>5	54	58.1	24	27.9
Sibling health problems				
No	45	83.3	31	77.5
Yes	9	16.7	9	22.5
Maternal education level				
High school or less	55	59.1	40	46.5
More than high school	38	40.9	46	53.5
Maternal employment status				
Unemployed	52	55.9	40	46.5
On leave	31	33.3	26	30.2
Working	10	10.8	20	23.3
Socioeconomic status[c]				
Category I (high)	7	7.6	15	17.4
Category II	19	20.4	11	12.8
Category III	19	20.4	20	23.3
Category IV	24	25.8	18	20.9
Category V (low)	24	25.8	22	25.6
Prior family problems				
None	26	28.0	21	24.4
Few	31	33.3	43	50.0
Some	30	32.2	21	24.4
Many	6	6.5	1	1.2

[a]Other than apnea, bradycardia, and monitoring.
[b]$\chi^2 = 16.528$, $df = 1$, $P < .0001$.
[c]Hollingshead Two-factor Scale.

TABLE 2—Means for Case and Comparison Mothers on Potentially Confounding Variables

	Case Mothers		Comparison Mothers	
	Mean	SD	Mean	SD
Infant gestational age at birth, mo	35.5	4.6	37.6*	3.0
Difficult temperament, mean score[a]	11.2	4.0	10.2	3.7
Maternal age, y	26.7	5.5	27.4	6.5
No. of children at home	1.8	0.8	1.6	0.8
Years married or living together	5.1	3.8	5.2	4.5
Satisfaction with social support, mean score[b]	54.1	7.4	56.3**	5.5

[a]Scale developed for this study; conceptually based on fussy-difficult scale of Bates' Infant Characteristics Questionnaire (1979), but modified (with permission) for use in telephone interviewing.
[b]Cmic's Social Support measure, satisfaction scale: a higher score implies greater satisfaction.
*$t = -3.6$, $P = .0004$.
**$t = -2.23$, $P = .0268$.

infant (alpha of 0.78) and parent–sibling (alpha of 0.85) relationships and ease of daily activities (alpha of 0.82). Preliminary reliability and validity testing of these scales supported their use.[2]

Scores on scales measuring dependent variables were dichotomized at predetermined cutoff points to distinguish normal or optimal functioning from at-risk or suboptimal functioning. We examined confounding separately for each dependent variable. We then used logistic regression to evaluate the relationship between case-comparison status and each dependent variable while controlling for confounders. We conducted separate analyses for mothers and fathers. SAS, version 5,[19] and PROC LOGIST[20] were used.

Results

At the time of the study, case infants had been monitored a mean of 45 days (median = 40, mode = 30). They were monitored for the following reasons: (1) SIDS sibling (5.4%); (2) severe apparent life-threatening event (20.4%); (3) less serious event or abnormal pneumogram or polysomnogram (full-term infants) (24.8%); or (4) less serious event, usually apnea of prematurity continuing up to approximately the time of discharge (preterm infants) (49.5%). Reason for monitoring, presence or absence of additional medical problems, gestational age, frequency of reported alarms, and frequency of need to stimulate the infant to breathe were not significantly associated with any of the dependent variables.[2] As a result, heterogeneity in these factors was considered a strength rather than a weakness of the study.

The mean age of the infants was 3.0 months. Most were White (69.9%) and male (58.1%). Most of the mothers were married (63.4%), had more than one child (55.4%), and had commercial insurance (63.0%).

Case and comparison groups differed on only three of many potentially confounding variables examined: gestational age, days hospitalized at birth, and satisfaction with social support (Tables 1 and 2).

Among mothers, gestational age confounded the relationship between case-comparison status and one dependent variable, relationship with baby. Satisfaction with social support was a confounder in relation to several of the dependent variables: parental health, parental sense of competence, parental role restriction, relationship with spouse, marital satisfaction, and social isolation. No other study variables were confounders among mothers, and none were confounders among fathers.

When confounders were controlled using logistic regression, case and comparison mothers differed significantly on only one dependent variable: parental health (Table 3). Case and comparison fathers did not differ significantly on any of the dependent variables (Table 4).

TABLE 3—Mothers: Beta Coefficients from Logistic Regression of Case-Comparison Group and Confounders on Dependent Variables

Dependent Variable[a]	Independent Variable(s)	β	SE	P
Parental depression	Group[b]	0.485	0.541	.3707
Poor parental health	Group	1.109	0.372	.0028[c]
	Satisfaction with support	−0.083	0.027	.0021
Sibling behavior problems	Group	1.609	0.819	.0504
Low parental sense of competence	Group	0.407	0.665	.5405
	Satisfaction with support	−0.048	0.040	.2354
Poor parental attachment	Group	−0.137	0.407	.7361
High parental role restriction	Group	0.556	0.438	.2036
	Satisfaction with support	−0.072	0.029	.0127
Poor relationship with spouse	Group	−0.079	0.503	.8751
	Satisfaction with support	−0.138	0.035	.0001
Low marital satisfaction	Group	0.660	0.383	.0852
	Satisfaction with support	−0.130	0.332	.0001
Poor relationship with baby	Group	−0.056	0.319	.8597
	Gestational age	−0.114	0.042	.0062
Poor relationship with sibling	Group	−0.470	0.520	.3662
High social isolation	Group	0.334	0.465	.4736
	Satisfaction with support	−0.133	0.033	<.0001
Low ease of daily activities	Group	0.437	0.313	.1629

Note. Confounding was controlled by entering variables into the regression model simultaneously.
[a]Reference category for dependent variables is normal or optimal functioning.
[b]Reference category is comparison group.
[c]Odds ratio = 3.03 (95% confidence interval = 1.46–6.30).

Discussion

The difference between the scores of case and comparison mothers on the Parental Health measure indicates a greater risk of poor health among mothers of monitored infants. We found no other significant differences between case and comparison families based on either maternal or paternal responses.

Poor health, fatigue, and somatic complaints have been reported among mothers of monitored infants.[6,9,14] While McElroy et al.[13] found no differences in somatization among case and comparison mothers, study infants were no longer being monitored at the time somatization was measured. In this study, only two of the five items on the Parental Health measure distinguished case vs comparison mothers, and these were global health measures rather than questions about specific symptoms (see Appendix). It has not been established whether global health measures are accurate indicators of physical health or are more precisely interpreted as measures of a sense of well-being.[21–25] Regardless, since case and comparison infants in this study differed in the need for home monitoring as well as in the use of monitors, and prior maternal health status was not measured, poor maternal health cannot be definitively attributed to the experience of home monitoring.

The absence of significant differences, other than maternal health risk, between case and comparison families indicates that the majority of families in supportive monitoring programs have adequate resources enabling them to cope well with the demands and stresses of home monitoring and to function capably in many aspects of family life. □

TABLE 4—Fathers: Beta Coefficients from Logistic Regression of Case-Comparison Group on Dependent Variables

Dependent Variable[a]	Independent Variable	β	SE	P
Parental depression	Group[b]	−0.06	1.02	.9526
Poor parental health	Group	−0.16	0.52	.7556
Sibling behavior problems	Group	−0.07	0.84	.9345
Low parental sense of competence	Group	0.53	0.89	.5511
Poor parental attachment	Group	−0.08	0.43	.8475
High parental role restriction	Group	−1.81	1.11	.1034
Poor relationship with spouse	Group	−0.47	0.79	.5547
Low marital satisfaction	Group	0.82	0.44	.0595
Poor relationship with baby	Group	1.06	0.64	.0944
Poor relationship with sibling	Group	−0.07	0.79	.9304
High social isolation	Group	−0.28	0.59	.6347
Low ease of daily activities	Group	−0.31	0.48	.5237

Note. There were no confounding variables among fathers.
[a]Reference category for dependent variables is normal or optimal functioning.
[b]Reference category is comparison group.

Acknowledgments

This research was supported in part by Grant HS 05988 from the National Center for Health Services Research and Health Care Technology Assessment and in part by a Scholl Fellowship from the National Sudden Infant Death Syndrome Foundation. This research was done in partial fulfillment of the requirements for the degree of Doctor of Science from the Department of Maternal and Child Health of The Johns Hopkins University School of Hygiene and Public Health.

Thanks to the members of the primary author's dissertation committee for their consultation at various phases of this research. Thanks also to staff at the SIDS Institute and the Department of Obstetrics and Gynecology, the University of Maryland Medical Center, Baltimore; the staff of the Apnea Clinic at Children's National Medical Center, Washington, DC; and the staff in the Neonatology and Medical Records Departments at Holy Cross Hospital, Silver Spring, Md, for their assistance.

References

1. US Congress, Office of Technology Assessment. *Technology-Dependent Children: Hospital v. Home Care—A Technical Memorandum.* Washington, DC: US Government Printing Office; 1987. OTA-TM-H-38.
2. Ahmann E. *Home Apnea Monitoring and Patterns of Family Life.* Baltimore, Md: The Johns Hopkins University School of Hygiene and Public Health; 1989. Dissertation.
3. American Academy of Pediatrics Task Force on Prolonged Apnea. Prolonged apnea. *Pediatrics.* 1978;61:651–652.
4. National Institutes of Health. *Infantile Apnea and Home Monitoring.* Washington, DC: US Department of Health and Human Services, Public Health Service; 1986. NIH publication no. 87-2905.
5. Nelson NM. But who shall monitor the monitor? *Pediatrics.* 1978;61:663–664.
6. Black L, Hersher L, Steinschneider A. Impact of the apnea monitor on family life. *Pediatrics.* 1978;62:681–685.
7. Cain LP, Kelly DH, Shannon DC. Parents' perceptions of the psychological and social impact of home monitoring. *Pediatrics.* 1980;66:37–41.
8. Fiser DH, Lyons L. Family acceptance of home apnea monitors. *Am J Med Sci.* 1985; 290:52–55.
9. Wasserman AL. A prospective study of the impact of home monitoring on the family. *Pediatrics.* 1984;74:323–329.
10. Nuttall P. Maternal responses to home apnea monitoring of infants. *Nurs Res.* 1988; 37:354–357.
11. Bendell RD, Culbertson JL, Shelton TL, et al. Interrupted infantile apnea: impact on early development, temperament, and maternal stress. *J Clin Child Psychol.* 1986;15:304–310.
12. Vohr BR, Chen A, Coll CG, et al. Mothers of preterm and full-term infants on home apnea monitors. *Am J Dis Child.* 1988; 142:229–231.
13. McElroy E, Steinschneider A, Weinstein S. Emotional and health impact of home monitoring on mothers: a controlled prospective study. *Pediatrics.* 1986;78:780–786.
14. Desmarez C, Blum D, Montauk L, et al. Impact of home monitoring for sudden infant death syndrome on family life. *Eur J Pediatr.* 1987;146:159–161.
15. Abidin RR. *Parenting Stress Index.* 2nd ed. Charlottesville, Va: Pediatric Psychology Press; 1986.
16. Nunnally JC. *Psychometric Theory.* 2nd ed. New York, NY: McGraw-Hill; 1978: 225–255.
17. Schumm WA, Nichols CW, Schectman KL, et al. Characteristics of responses to the Kansas Marital Satisfaction Scale by a sample of 84 married mothers. *Psychol Rep.* 1983;53:567–572.
18. Eyberg SM, Ross AW. Assessment of child behavior problems: the validation of a new inventory. *J Clin Child Psychol.* 1978; 7:113–116.
19. SAS Institute Inc. *SAS, Version 5.* Cary, NC: SAS Institute Inc; 1985.
20. Harrell FE Jr. *PROC LOGIST.* The SUGI Supplemental Library User's Guide, Version 5. Cary, NC: SAS Institute Inc., 1986.
21. Garrity TF, Somes GW, Marx MB. Factors influencing self assessment of health. *Soc Sci Med.* 1978;12:77–81.
22. Ware JE. The assessment of health status. In: Aiken LA, Mechanic D, eds. *Applications of Social Science to Clinical Medicine and Health Policy.* New Brunswick, NJ: Rutgers; 1986:204–228.
23. Mechanic D, Hansell S. Adolescent competence, psychological well being, and self-assessed physical health. *J Health Soc Behav.* 1987;28:364–374.
24. Tessler R, Mechanic D. Psychological distress and perceived health status. *J Health Soc Behav.* 1978;19:254–262.
25. Mechanic D. Discussion of research programs on relations between stressful life events and episodes of physical illness. In: Dohrenwend BS, Dohrenwend BP, eds. *Stressful Life Events: Their Nature and Effects.* New York, NY: Wiley; 1974:87–97.

APPENDIX—Items 98 and 101 of the Parenting Stress Index[15] (Parental Health Subscale)

Physically I feel good most of the time.
1. Strongly agree
2. Agree
3. Not sure
4. Disagree
5. Strongly disagree

Since I've had my child:
1. I have been sick a great deal.
2. I haven't felt as good.
3. I haven't noticed any change in my health.
4. I have been healthier.

Carrier Screening for Cystic Fibrosis: A Pilot Study of the Attitudes of Pregnant Women

Jeffrey R. Botkin, MD, MPH, and Sonia Alemagno, PhD

ABSTRACT

We surveyed the attitudes of a consecutive sample of 306 pregnant Caucasian women toward carrier screening for cystic fibrosis. Of the 214 respondents, 98% said that screening should be offered before pregnancy, and 69% said they would accept carrier screening during pregnancy. Twenty-nine percent of the respondents indicated a willingness to terminate a pregnancy if the fetus were found to have cystic fibrosis. We conclude that carrier screening is of interest to pregnant women, although interest in terminating a pregnancy because of screening results may be limited. (*Am J Public Health.* 1992;82:723–725)

Introduction

The identification of the gene responsible for cystic fibrosis (CF) offers the capability of performing carrier screening for CF in the general population. Cystic fibrosis is an autosomal recessive disease and is the most common lethal genetic disease in Caucasians. The incidence at birth is approximately 1 per 2500 Caucasian infants. The carrier (heterozygote) frequency is approximately 1 in 20 to 25 Caucasians in North America.[1] Carriers for CF are asymptomatic; however, individuals may be interested in knowing their genetic status in order to make informed reproductive decisions.

Carrier screening for CF in the general population is controversial.[2–5] The initiation of carrier screening in the general population has been inhibited by the relatively low sensitivity of the current technology for screening[6–8]; to date, only 80% to 85% of the CF mutations have been identified in most populations.[9]

If carrier screening for CF is offered to the general population, a primary interest group is likely to be pregnant women. A body of literature exists that describes and analyzes the attitudes and behavior of women toward screening and prenatal diagnosis in a variety of circumstances.[10–18] However, CF has its own distinctive features as a disease, including its morbidity and mortality, risk levels, and the population group at risk. This pilot study was undertaken to explore the interest in carrier screening for CF in pregnant women without a family history of CF and to determine the attitudes of such women toward the potential use of this information in making reproductive decisions.

Methods

Subjects were recruited from the obstetric clinics at a large urban county hospital in Cleveland and at a suburban health maintenance organization in northeastern Ohio. A consecutive sample of 306 pregnant Caucasian women at less than 18 weeks' gestation was recruited to participate in the study. The participants' baseline knowledge of CF was assessed by their responses to three questions; the women were then asked to read educational materials and complete a 40-item questionnaire at home. The participants were encouraged to complete the questionnaire with the father of the baby, when appropriate. Respondent confidentiality was maintained. Pretest responses were available from 84 women who did not return the questionnaire.

Part 1 of the questionnaire assessed demographic information, attitudes about the present pregnancy, and history of previous pregnancies. Part 2 of the questionnaire consisted of information about CF, the carrier state, and reproductive options for carriers. The information was written at a 10th-grade reading level. The information was reviewed for accuracy and balance by parents of persons with CF and by medical specialists. Part 3 of the questionnaire consisted of a test of factual knowledge about CF, followed by items designed to ascertain attitudes about testing and reproductive decisions. Most questions were designed to elicit yes or no responses on a 4-point scale: (1) Yes, definitely; (2) Yes, probably; (3) No, probably not; (4) No, definitely not.

Data were analyzed with the Statistical Package for the Social Sciences/PC+. Analysis for this project consisted primarily of univariate and bivariate analysis, including cross-tabulation and correlational procedures.

Results

Of the 306 questionnaires distributed, 214 were returned in usable form, for a

Jeffrey R. Botkin is with the Department of Pediatrics, Center for Biomedical Ethics and Sonia Alemagno is with the Department of Family Medicine at Case Western Reserve University School of Medicine, Cleveland, Ohio.

Requests for reprints should be sent to Jeffrey R. Botkin, MD, MPH, Department of Pediatrics, MetroHealth Medical Center, 3395 Scranton Road, Cleveland, OH 44109.

This paper was submitted to the Journal July 2, 1991, and accepted with revisions December 9, 1991.

TABLE 1—Demographic Characteristics of Respondents (n = 214)

Age, y	27.0 ± 5.7 (range 15–42)
Education,	%
High school or less	42
Technical school	3
Some college	22
College graduate	25
Advanced degree	7
Employment	
Homemakers	31
Laid off	1
Unemployed	7
Students	7
Full-time employment	40
Part-time employment	14
Religion	
No affiliation	28
Catholic	40
Jewish	1
Pentecostal/Southern Baptist	7
Other (Protestant denom.)	24
Marital Status	
Married	81
Parity	
0	42
1	36
2	14
3+	8
Previous abortions (n = 145)	
Yes	25
No	75

TABLE 2—Posteducation Knowledge of Cystic Fibrosis (CF) (n = 214)

Question	Correct Response (% Who Answered Correctly)
1. CF is a disease of the ____?	Lungs and digestion (100)
2. Children with CF are usually ____?	Not retarded (99)
3. Most people with CF will die by ____?	30 years of age (93)
4. Carriers for CF show no signs of the disease	True (96)
5. Chance of CF in a child of two carriers is ____?	1 in 4 (93)
6. Chance *you* are a carrier for CF?	1 in 20 (53)
7. Brothers and sisters of carriers are likely to be carriers.	True (87)

TABLE 3—Participants' Interest in Carrier Screening for Cystic Fibrosis (n = 214)

Question	Responses, %			
	Yes (Total)[a]	Yes, definitely	No (Total)[b]	No, definitely not
1. Is it important to know carrier status before marriage?	75	33	25	6
2. Should carrier screening be offered before pregnancy?	98	69	2	1
3. Would *you* have taken the test before pregnancy?	84	58	16	2
4. Would *father* have taken the test before pregnancy?	79	48	21	4
5. Would *you* have the test done now during pregnancy?	69	40	31	14
6. Would *father* have the test done now during pregnancy?	63	34	37	13

[a]Yes (Total) = "Yes, definitely" + "Yes, probably."
[b]No (Total) = "No, definitely not" + "No, probably not."

response rate of 70%. There were no variables with more than 2% missing data. The demographic characteristics of the study population (Table 1) demonstrate participation by women with a broad range of demographic and socioeconomic characteristics.

Seventy-eight percent of the respondents had heard of CF, and 14% of those had known someone with the disease, although none had siblings or children with CF. Of those who had heard of CF, 47% were aware that it was a disease of the lungs. There was no significant difference in baseline knowledge between responders and nonresponders (n = 84).

Responses to the knowledge questions (Table 2) demonstrate that the participants read the information and that the information was effective in transmitting factual knowledge of CF. It is notable that 11% of the women responded incorrectly that the risk of two carriers' producing an affected child is 100%. Respondents also were likely to underestimate their personal risk of being a carrier: 43% responded that they were at a 1 in 400 risk of being a carrier (the lowest risk level offered as a choice). For each of these two questions those answering incorrectly were less likely to have attained a high level of formal education than were those who answered correctly ($P \leq .01$).

The results shown in Table 3 demonstrate that the participants had a strong interest in being tested for CF carrier status before pregnancy (84%) and during pregnancy (69%). Those who had heard of CF ($P \leq .01$), those who knew someone with CF ($P \leq .01$), those who answered correctly on the nature of CF on the pretest ($P \leq .05$), and those who attended religious services frequently ($P \leq .01$) were *less* likely to want carrier testing during pregnancy. Age, education, and employment were not predictive of the desire to be tested, and those who underestimated their personal risk of carrier status were not less likely to want testing.

With respect to reproductive choices (Table 4), the majority of women (67%) said they would be interested in prenatal diagnosis if they were at risk for having a child with CF; however, a minority (29%) expressed a willingness to terminate a pregnancy if the fetus were found to have CF. Choices about abortion were not related to age, education, employment, pretest knowledge of CF, or perceived risk. Our sample contained a high proportion of Catholic women; however, Catholics were not less likely to choose abortion, nor were those who said they had a religious affiliation versus those who did not.

To estimate the proportion of affected pregnancies that would have been identified and terminated in our sample population, we analyzed the data to estimate the sample's use of testing for each stage of a screening protocol. These results (Table 5) suggest that approximately 24% of affected pregnancies among the women offered screening would have been identified and terminated as a result of a carrier screening program. This estimate assumes a 100% test sensitivity.

With respect to issues of confidentiality and free choice, 99% of the respondents indicated that they would be willing

TABLE 4—Participants' Interest in Using Cystic Fibrosis (CF) Screening to Make Reproductive Choices (n = 214)

Question	Responses, %			
	Yes (Total)[a]	Yes, definitely	No (Total)[b]	No, definitely not
1. Is carrier status a good reason for two carriers not to marry?	9	0	91	43
2. If *you* and the *father* were both carriers, would you want the baby tested in the womb?	67	44	33	16
3. If you found your child would have CF, would you have an abortion?	29	9	71	42

[a]Yes (Total) = "Yes, definitely" + "Yes, probably."
[b]No (Total) = "No, definitely not" + "No, probably not."

TABLE 5—Estimated Sequential Use of Services

A. Proportion of pregnant women accepting screening	69%
B. Proportion of (A) with father accepting screening	86%
C. Proportion of (A+B) choosing prenatal diagnosis	86%
D. Proportion of (A+B+C) choosing termination	47%
Complete utilization = A × B × C × D = 24%	

to share their test results with family members. A large majority (80%) felt that carrier testing should not be required by law, and 60% said that physicians should not be required to divulge the results of CF testing to family members against the wishes of the patient.

Discussion

The results demonstrate a substantial interest in carrier screening in this sample and support a design for a program that is voluntary and confidential. Although the majority of the women were interested in knowing their carrier status, most indicated a reluctance to use the information to prevent the birth of an affected child. Of particular interest is our finding that those who were more familiar with CF tended to be less interested in carrier screening. The hypothetical nature of this study limits our ability to make confident predictions of actual behavior in a screening program, and attitudes may vary with time, clinical approach, and test validity. □

Acknowledgments

This work was supported by a grant from the Ohio Board of Regents.

The results of this study were presented, in part, at the national meeting of the American Association for the Advancement of Science, Washington, DC, February 2, 1991, and at the meeting of the Society for Pediatric Research, New Orleans, La, May 1, 1991.

We would like to thank Rose Harcar, RN, Sue Fitz, RN, Linda Cruise, RN, Graham Ashmead, MD, and Ellen Lazarus, PhD, for their assistance with this research and Ruth Faden, PhD, Tom Murray, PhD, Irwin Schafer, MD, and Stephen Zyzanski, PhD, for review of the manuscript.

References

1. Boat TF, Welsh MJ, Beaudet AL. Cystic fibrosis. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic Basis of Inherited Disease.* 6th ed. New York: McGraw Hill; 1989:2649–2680.
2. Wilfond BS, Fost N. The cystic fibrosis gene: medical and social implications of heterozygote detection. *JAMA.* 1990;263: 2777–2783.
3. Colten HR. Screening for cystic fibrosis: public policy and personal choices. *N Engl J Med.* 1990;322:328–329.
4. Roberts L. To test or not to test. *Science.* 1990;247:17–19.
5. Beaudet AL. Carrier screening for cystic fibrosis. *Am J Hum Genet.* 1990;47:603–605.
6. Roberts L. CF gene proves uncooperative. *Science.* 1990;250:1077.
7. Caskey CT, Kaback M, Beaudet AL, et al. The American Society of Human Genetics statement on cystic fibrosis screening. *Am J Hum Genet.* 1990;46:393.
8. Statement from the National Institutes of Health Workshop on Population Screening for the Cystic Fibrosis Gene. *N Engl J Med.* 1990;323:70–71.
9. Zielenski J, Markiewicz D, Rininsland F, Rommens J, Tsui LC. A cluster of highly polymorthic dinucleotide repeats in intron 17b of the cystic fibrosis transmembrane conductance regulator (CFTR) gene. *Am J Hum Genet.* 1991;49:1256–1262.
10. Faden RR, Chwalow AJ, Quaid K, et al. Prenatal screening and pregnant women's attitudes toward the abortion of defective fetuses. *Am J Public Health.* 1987;77:288–290.
11. Sell RR, Roughmann KJ, Doherty RA. Attitudes toward abortion and prenatal diagnosis of fetal abnormalities: implications for educational programs. *Soc Biol.* 1978; 25:288–301.
12. Case GA, Faden RR, Holtzman NA, et al. Assessment of risk by pregnant women: implications for genetic counseling and education. *Soc Biol.* 1986;33:57–64.
13. Roughmann KJ, Doherty R, Robinson JL, et al. The selective utilization of prenatal genetic diagnosis. *Medical Care.* 1983;21: 1111–1125.
14. Adams MM, Finley S, Hansen H, et al. Utilization of prenatal genetic diagnosis in women 35 years of age and older in the United States, 1977 to 1978. *Am J Obstet Gynecol.* 1981;139:673–677.
15. Marion JP, Kassam G, Fernhoff PM, et al. Acceptance of amniocentesis by low-income patients in an urban hospital. *Am J Obstet Gynecol.* 1980;138:11–15.
16. Modell B, Ward RH, Fairweather DV. Effect of introducing antenatal diagnosis on reproductive behavior of families at risk for thalassemia major. *Br Med J.* 1980;280: 1347–1350.
17. Goldstein MS, Greenwald S, Nathan T, et al. Health behavior and genetic screening for carriers of Tay-Sachs disease: a prospective study. *Soc Sci Med.* 1977;11:515–520.
18. Beck E, Blaichman S, Scriver CR, Clow CL. Advocacy and compliance in genetic screening: behavior of physicians and clients in a voluntary program of testing for the Tay-Sachs gene. *N Engl J Med.* 1974;291:1166–1170.

A Cohort Study of the Impact of Perinatal Drug Use on Prematurity in an Inner-City Population

Joseph G. Feldman, DrPH, Howard L. Minkoff, MD, Sandra McCalla, MD, and Martin Salwen, MD

ABSTRACT

Perinatal illicit drug use was determined by anonymously testing the urine of 1111 inner-city parturients. Fourteen percent tested positive, 11.5% for cocaine. As determined by self-report or by positive test results, 17% used drugs. Twenty-eight percent of low birthweights and 20% of premature births in this group were attributable to drug use. Risk of prematurity was reduced by a factor of 2 to 3 with adequate prenatal care. These data suggest an enormous potential for improving perinatal health. (*Am J Public Health.* 1992;82:726–728)

Introduction

Cocaine use has increased substantially among pregnant women over the past 10 years,[1,2] in association with adverse effects such as low birthweight, prematurity, and small head circumference.[3–6]

Most previous studies used small and selective samples.[3,6] We studied the urine of 1111 inner-city women who delivered over a 4½ month period, and found that cocaine metabolites in the urine of parturients were associated with decreased use of prenatal care services, higher parity, more sexually transmitted diseases, lower birthweight, and more preterm births.[7,8] In the present study, we estimate the impact of drug use on low birthweight and prematurity in an inner-city community.

Methods

Our study methods have been described previously.[7] Women delivering at a municipal hospital in New York City between October 18, 1988 and March 1, 1989 (1331 women) were eligible for participation. Entry criteria were admission and/or delivery on a shift staffed by study personnel. Specimens obtained from parturients for routine urinalysis were saved, refrigerated, given code numbers, and results combined with the other patient information.

Data collected included demographic information, history of drug use, and pregnancy outcome measures, which were available for 1111 patients. Study pediatricians assessed all newborns during several shifts per week but were not always available, and so the number of infants included in the study (846) was less than the number of mothers. Pediatric assessment included head circumference, length, and gestational age (judged on the basis of a physical examination).

Prenatal care was categorized as received at the delivering institution (YES); received at a different institution at least once during pregnancy (OTHER); no care reported (NONE).

Urine screening for illicit drugs was done on a Hitachi 705 analyzer with a homogeneous enzyme-multiplied immunoassay technique[9–10] (details available on request).

Statistical Methods

Determination of exposure to drugs, either by urine toxicology or by history, was made before pregnancy termination and without knowledge of the birth outcome.

Multiple logistic regression models were used to estimate the association of drug exposure to birth outcome, adjusting for other factors including cigarette use, alcohol use, hypertension, maternal age, parity, and use of prenatal care.[11]

The population attributable risk (AR), the proportion of prematurity in the community that was caused by the mothers' exposure to drugs during pregnancy, was estimated as follows[12]:

$$AR = \frac{(\%\text{ of drug use}) \times (\text{relative risk of outcome} - 1)}{1 + (\%\text{ of drug use}) \times (\text{relative risk of outcome} - 1)}$$

Relative risks were estimated from multiple logistic regression models, adjusting for the variables noted above.

Results

The 265 mothers for whom information about offspring was unavailable were

Joseph G. Feldman is with the Department of Preventive Medicine, Howard L. Minkoff and Sandra McCalla are with the Department of Obstetrics and Gynecology, and Martin Salwen is with the Department of Pathology, all at the State University of New York, Health Science Center at Brooklyn.

Requests for reprints should be sent to Joseph G. Feldman, DrPH, Department of Preventive Medicine, SUNY, Health Science Center at Brooklyn, 450 Clarkson Avenue, Box 43, Brooklyn, NY 11203-2098.

This paper was submitted to the Journal June 25, 1991, and accepted with revisions November 12, 1991.

similar to the 846 for whom pediatric information was available with respect to drug use and demographic characteristics.

The average age was 25.6 ± 6.0 years. Ninety-one percent of the parturients were Black (two thirds were born in the United States), and 9% were of Hispanic origin. Median parity was two. About two thirds were unemployed. Of drug users, 11.5% used cocaine, 1.2% used marijuana, 1.3% used opiates, 0.4% used methadone, and 3.1% had negative urine tests but positive histories. Table 1 shows the regression coefficients, *P* values, and the odds ratios and their 95% confidence intervals for the variables included in the final model for predicting birthweight of less than 2500 g. Drug users were at 3.3 times greater risk of giving birth to a child weighing less than 2500 g, independent of other factors, such as age, parity, hypertension, alcohol, or prenatal care. Agreement was good between the observed and expected frequencies of low birthweight by deciles of risk, based on the model (Chi-square [8 *df*] = 7.58, *P* = .48). Table 2 shows the results for the logistic regression with premature birth as the outcome. The impact of the predictor variables was similar for both birthweight and premature birth. Statistically significant predictors in both models were hypertension, prenatal drug use, and absence of prenatal care.

Table 3 shows observed and predicted odds ratios for low birthweight and premature birth categories of drug use and prenatal care. There was good agreement between the observed and predicted odds ratios, again suggesting that the logistic regression models were representing the data adequately. Drug use and absence of prenatal care were each associated with a two- to threefold elevation in the risk of premature birth. Absence of prenatal care combined with drug use increased risk six- to ninefold.

Table 4 shows the actual percentages of low birthweight and premature birth for the categories of prenatal care and drug use.

Using the definition of attributable risk described earlier, and standardized odds ratios of 3.34:1 for low birthweight and 2.64:1 for premature birth in infants exposed, estimates of community attributable risk due to drug use were as follows:

$$\text{AR (low birthweight)} = \frac{.171 \times 2.24}{1 + (.171 \times 2.24)} = 27.7\%$$

TABLE 1—Multiple Logistic Regression Analysis of Outcome Birthweight Less than 2500 G

Variable	Regression Coefficient	*P* Value	Odds Ratio	95% Confidence Interval
Age	−.0001	.970	1.00	.96–1.04
Cigarettes (YES vs NO)	.140	.617	1.15	.67–1.99
Parity	0.53	.534	1.06	.89–1.24
Hypertension (YES vs NO)	1.548	.000	4.70	2.33–9.49
Alcohol (YES vs NO)	.041	.904	1.04	.53–2.03
Drug Use (YES vs NO)	1.206	.000	3.34	1.95–5.72
Prenatal Care[a] (NO vs OTHER)	.173	.375	1.19	.65–2.18
Prenatal Care[b] (YES vs OTHER)	−.999	.000	.37	.21–11

[a]No prenatal care vs care at another institution.
[b]Prenatal care at delivering institution vs another institution.

TABLE 2—Multiple Logistic Regression Analysis of Outcome Gestation Less than 37 Weeks

Variable	Regression Coefficient	*P* Value	Odds Ratio	95% Confidence Interval
Age	−.026	.380	.98	.92–1.03
Cigarettes (YES vs NO)	.431	.264	1.54	.72–3.28
Parity	.164	.144	1.18	.95–1.47
Hypertension (YES vs NO)	1.673	.000	5.33	1.99–11
Alcohol (YES vs NO)	−.137	.763	.87	.36–2.12
Drug Use (YES vs NO)	.970	.015	2.64	1.21–5.78
Prenatal Care[a] (NO vs OTHER)	−.032	.938	.97	.43–2.19
Prenatal Care[b] (YES vs OTHER)	−.826	.021	.44	.22–11

[a]No prenatal care vs care at another institution.
[b]Prenatal care at delivering institution vs another institution.

TABLE 3—Observed and Predicted Odds Ratios of Low Birthweight and Gestation <37 Weeks by Prenatal Drug Use and Absence of Prenatal Care

Prenatal Drug Use	Prenatal Care	Birthweight <2500 g		Gestation <37 wk	
		Observed	Predicted	Observed	Predicted
Yes	DI	3.2	3.3	3.8	2.7
	Other	14.0	9.1	10.9	5.8
	None	8.1	10.8	4.2	6.0
No	DI	1.0	1.0	1.0	1.0
	Other	1.7	2.7	1.5	2.2
	None	2.9	3.2	3.6	2.3

Note. DI = delivering institution.

TABLE 4—Frequency of Low Birthweight (<2500 g) and Gestation <37 Weeks by Prenatal Care and Drug use

Drug Use	Total	Birthweight <2500 g, %	Total	Gestation <37 wk, %
		Care at Delivering Institution		
Yes	72	23.6	48	20.8
No	723	8.9	513	6.4
		Care at Another Institution		
Yes	33	57.6	28	42.9
No	106	14.2	87	9.2
		No Prenatal Care		
Yes	77	44.2	49	22.5
No	73	21.9	45	20.0

$$AR\ (prematurity) = \frac{.154 \times 1.64}{1 + (.154 \times 1.64)} = 20.2\%$$

Discussion

In an inner-city hospital, 17% of parturients had evidence of drug use during pregnancy, with cocaine accounting for 83% of positive toxicologies.

After adjustment for several factors such as hypertension, age, parity, and cigarettes, drug use was estimated to be responsible for between 20% and 30% of the premature births occurring in the parturients' community. Documented prenatal care was associated with a two- to threefold reduction in prematurity, even in drug users.

Prematurity is a preeminent public health problem in the United States,[13-14] and cocaine use during pregnancy has been associated with preterm births.[1,4,5,7,15] Although a physiologic basis for cocaine-derived preterm birth has been suggested,[16] some authors have not found an increase in population prematurity rates despite rapid increases in rates of cocaine use.[17] Our data support not merely an association between cocaine and preterm birth, but a substantive contribution of cocaine to prematurity within an inner-city community.

The relationship among cocaine use, the absence of prenatal care, and poor pregnancy outcomes[7,8] partly reflects the more desperate circumstances and more intense drug use of "unregistered" patients. Most drug users (70%) were also cigarette smokers, and cigarette smoking is known to cause prematurity. It is probable that some portion of the 20% to 30% risk of prematurity attributable to drug use in this study would remain if parturients persisted in smoking during pregnancy. Nonetheless, one third of low-birthweight births in the United States occur in populations similar to the one studied (poor and minority), suggesting that drug use, particularly use of cocaine, contributes substantially to national infant morbidity. □

References

1. Abelson HI, Miller JD. A decade of trends in cocaine use in the household population. *NIDA Res Monogr.* 1985;61:35–49.
2. Krayewuski KJ. Crack and cocaine—current medical issues. *Texas Med.* 1988; 84:48–59.
3. Chasnoff IJ, Burnsett WAJ, Schnoll SH, Burns KA. *Cocaine use in pregnancy. N Engl J Med.* 1985;313:666–669.
4. Cherukuri R, Minkoff HL, Feldman JG, et al. A cohort study of alkaloidal cocaine ("crack") in pregnancy. *Obstet Gynecol.* 1988;72:147.
5. Zuckerman B, Frank D, Hingson R, et al. Effects of maternal marijuana and cocaine use on fetal growth. *N Engl J Med.* 1989;320:762–768.
6. Neerhof MG, MacGregor SN, Retsky SS, Sullivan TP. Cocaine abuse during pregnancy: peripartum prevalence and perinatal outcome. *Am J Obstet Gynecol.* 1989;161:633–638.
7. McCalla S, Minkoff HL, Feldman JG, et al. The biologic and social consequences of perinatal cocaine use in an inner-city population: results of an anonymous cross-sectional study. *Am J Obstet Gynecol.* 1991;164:625–630.
8. Minkoff HL, McCalla S, Delke I, et al. The relationship of HIV and syphilis to cocaine use. *Am J Obstet Gynecol.* 1990;163:521–525.
9. Baselt RC. Urine drug screening by immunoassay: interpretation of results. In: Baselt RC, ed. *Advances in Analytical Toxicology.* Foster City, Calif: Biomedical Publications; 1984;1:88–123.
10. Schwarz RH, Hawks RL. Laboratory detection of marijuana use. *JAMA.* 1985;245: 788–792.
11. Hosmer D, Lemeshow S. *Applied Logistic Regression.* New York, NY: John Wiley & Sons; 1989: chap 5.
12. Kahn H, Sempos C. *Statistical Methods in Epidemiology.* New York, NY: Oxford University Press; 1989: chap 4.
13. Fuchs F. Prevention of prematurity. *Am J Obstet Gynecol.* 1976;126:809–820.
14. Escobedo MB. Follow-up of prematurely born infants. *Clin Obstet Gynecol.* 1988; 31:662–667.
15. Chasnoff IJ, Griffith DR, MacGregor S, et al. Temporal patterns of cocaine use in pregnancy: perinatal outcome. *JAMA.* 1989;261:1741–1744.
16. Woods JR, Jr, Plessinger MA, Clark KE. Effect of cocaine on uterine blood flow and fetal oxygenation. *JAMA.* 1987;257:957–961.
17. Ney JA, Dooley SL, Keith LG, Chasnoff IJ, Socol ML. The prevalence of substance abuse in patients with suspected preterm labor. *Am J Obstet Gynecol.* 1990;162: 1562–1571.

The Impact of Mailing Psychoeducational Materials to Women with Abnormal Mammograms

Caryn Lerman, PhD, Eric Ross, ScM, Alice Boyce, MA, Pat McGovern Gorchov, MPH, Robert McLaughlin, Barbara Rimer, DrPH, and Paul Engstrom, MD

ABSTRACT

A randomized trial was conducted to evaluate the impact of mailed psychoeducational materials on adherence to subsequent annual mammography among women with prior abnormal mammograms. The results showed a 13% increment in adherence among women who received this intervention. This effect was independent of all sociodemographic and medical variables examined. We conclude that mailed psychoeducational materials may be an effective mechanism to improve adherence among women with abnormal mammogram results. (*Am J Public Health.* 1992;82:729–730)

Introduction

As use of mammography screening increases, an increasing number of women will have abnormal mammograms. The proportion of mammograms interpreted as abnormal in large screening programs is as high as 15% to 20%.[1,2] Thus, if 15% of the 48 million American women 40 years of age or older have mammograms, there would be more than 1 million abnormal mammograms per year.[3] Adherence to annual mammography screening is important for women with abnormal mammograms, because some of these women have an elevated risk for breast cancer.[4]

We conducted a randomized trial involving a mailed psychoeducational booklet to improve adherence to subsequent annual mammography among women with abnormal mammograms. The impact on adherence of two different styles of presenting educational information also was tested. On the basis of Prospect Theory,[5,6] we predicted that communications that emphasize the negative consequences of nonadherence (negative framing) would have a greater positive impact on adherence than equivalent communications that emphasize the positive consequences (positive framing).

Methods

Study Population

Respondents were 446 women, aged 50 to 74 years, who were members of the Health Maintenance Organization of Pennsylvania and New Jersey (HMO PA/NJ). Eligible respondents were those who had received an abnormal mammogram during the previous year and were eligible to receive an annual screening mammogram during the study period (March 1990 to May 1991). As part of their HMO membership, all respondents received an informational brochure and a referral for a free mammogram annually.

The study sample was selected from a pool of 936 women who met the above criteria. One hundred ninety-one women were excluded in the course of the study for the following reasons: discontinued HMO (10%), age ineligible (6%), cancer diagnosis (3%), and not eligible for mammogram (1%). Other women (n = 210) were excluded during the baseline survey phase because of unavailable telephone numbers (25%), refusals (3%), and language barriers (0.5%). An additional 89 women (17%) were lost during the follow-up survey phase for one of the above reasons. The final sample (n = 446) was not significantly different from the original sample in age, education, or mammogram results.

Procedures

Respondents were randomized into one of four study groups: (1) no survey—control, (2) survey—control, (3) survey—psychoeducational booklet/negative framing, and (4) survey—psychoeducational booklet/positive framing. The randomization was stratified by prior mammogram result (i.e., low suspicion/showing asymmetric breast tissue vs high suspicion/showing mass). Baseline telephone surveys of sociodemographic and medical characteristics were conducted with respondents randomized to groups 2, 3, and 4. Response rates across these three groups were not significantly different (overall rate = 71%).

One month later, intervention group respondents were mailed a 12-page psychoeducational booklet that described the meaning of abnormal mammograms and emphasized the necessity of continued screening. Suggestions for managing mammography-related anxiety also were included. Two versions of these printed

The authors are with the Fox Chase Cancer Center, Cheltenham, Pa.

Requests for reprints should be sent to Caryn Lerman, PhD, Fox Chase Cancer Center, 510 Township Line Road, Cheltenham, PA 19012.

This paper was submitted to the Journal August 14, 1991, and accepted with revisions December 31, 1991.

TABLE 1—Number and Percentage of Respondents Who Adhered to Mammography (n = 446)

Study Group[a]	Mammogram Result: Low Suspicion No.	%	High Suspicion No.	%	All Results No.	%
Control—no survey	49	49	31	60	80	53
Control—survey	32	59	18	49	50	55
Experimental—positive	40	71	22	59	62	66
Experimental—negative	46	65	27	69	73	67
All groups	167	59	98	59	265	59

[a]Study group effect significant ($\chi^2 = 7.76$, $P < .005$).

materials varied in terms of the "framing" of educational messages; however, the message content was equivalent.

One week later, all respondents were mailed the standard HMO breast screening packet, which included the free mammogram referral. Follow-up telephone calls were conducted with all respondents 3 months later, following expiration of the mammogram referrals, to assess adherence (response rate = 89%). Self-reported mammography adherence in this population has been shown to be highly reliable.[7]

Results

The average age of respondents was 60 years (SD = 6.8). The majority of women were White (85%) and married (71%). Eighty-two percent had at least a high school education. Sixty-three percent had a low-suspicion abnormal mammogram, and 37% had a high-suspicion abnormal mammogram. Fifteen percent had a biopsy after their abnormal mammogram; 16% received additional imaging (ultrasound or repeat mammogram) and 68% did not require diagnostic follow-up. There were no significant differences in these variables between study groups.

Overall, 53% of women in the two control conditions obtained their annual mammograms, compared with 66% of women who received the psychoeducational booklet ($\chi^2 = 7.76$, $P < .005$). Adherence rates in the two control groups were not different (53% vs 55%), suggesting that the baseline survey did not boost adherence. Adherence rates in the two intervention groups were not significantly different (66% vs 67%), suggesting that the framing of information had no effect.

Logistic regression analysis examined the effect of study group while controlling for the index of suspicion on the prior mammogram result and for sociodemographic variables. Study group (control vs booklet) was a significant independent predictor of subsequent mammography adherence (odds ratio = 1.7, $P < .005$). The interaction between study group and level of suspicion of the prior mammogram was not significant. No sociodemographic factor predicted adherence significantly.

Discussion

This randomized trial demonstrated a significant positive impact of a mailed psychoeducational booklet on mammography adherence among women with prior abnormal mammograms. The intervention groups demonstrated 13% greater adherence to subsequent annual mammography. Further analyses indicated that the framing of educational messages did not differentially affect adherence. Also, adherence was not associated with sociodemographic factors or prior mammogram result.

These results have implications for the management of women with abnormal mammograms. Our previous research has indicated that a substantial proportion of these women have psychological difficulties, even after learning that they do not have cancer.[1] Psychoeducational interventions tailored to the unique needs of these women potentially can ameliorate excessive distress and promote continued adherence.[8] Such communications should provide messages to heighten perceived vulnerability to breast cancer[9] but should be balanced with reassuring messages that emphasize the potential for early breast cancer detection and cure.[9–11]

This study was conducted in an HMO population. Women who elect to join an HMO may differ from the general population. However, our previous research suggests that members of HMO PA/NJ share the same sociodemographic backgrounds as nonmembers in the same geographic areas.[12] Future research is necessary to evaluate the impact of mailed educational materials in other populations with cost and access barriers. □

Acknowledgments

This research was supported by institutional grants awarded to the Fox Chase Career Center (National Institutes of Health Grants, CA 06927, RO3 CA 51691-01, and RR05895) and an appropriation from the Commonwealth of Pennsylvania.

References

1. Lerman C, Trock B, Rimer BK, Boyce A, Jepson C, Engstrom PF. Psychological and behavioral implications of abnormal mammograms. *Ann Intern Med.* 1991;114:657–661.
2. Winchester DP, Lasky HJ, Sylvester J, Maher ML. A television-promoted mammography screening pilot project in the Chicago metropolitan area. *CA.* 1988;38:291–309.
3. Dawson DA, Thompson DB, eds. *Breast Cancer Risk Factors and Screening: United States, 1987.* Hyattsville, Md: US Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Health Statistics; 1990; DHHS publication no. PHS 90-1500.
4. Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. *N Engl J Med.* 1985;312:146–151.
5. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science.* 1981;211:453–458.
6. Meyerowitz BE, Chaiken S. The effect of message framing on breast self-examination attitudes, intentions, and behavior. *J Pers Soc Psychol.* 1987;52:500–510.
7. King E, Rimer BK, Trock B, Balshem A, Engstrom P. How valid are mammography self-reports? *Am J Public Health.* 1990;80:1386–1388.
8. Lerman C, Rimer BK, Engstrom PF. Cancer risk notification: psychosocial and ethical implications. *J Clin Oncol.* 1991;9:1275–1281.
9. Lerman C, Rimer BK, Engstrom PF. Reducing avoidable cancer mortality through prevention and early detection regimens. *Cancer Res.* 1989;49:4955–4962.
10. Rimer BK, Keintz MK, Kessler HB, Engstrom PE, Rosan J. Why women resist mammograms: understanding patient-related barriers to acceptance of screening mammography. *Radiology.* 1989;172:243–246.
11. Paskett ED, White E, Carter WB, Chu J. Improving follow-up after an abnormal Pap smear: a randomized controlled trial. *Prev Med.* 1990;19:630–641.
12. Lerman C, Rimer B, Trock B, Engstrom P. Factors associated with repeat adherence to breast cancer screening. *Prev Med.* 1990;19:279–290.

Breast-feeding Patterns in a Rural Village in Giza, Egypt

Iman A. Hakim, PhD, and Iman A. El-Ashmawy, PhD

ABSTRACT

In a longitudinal study of infant feeding in rural Giza, Egypt, we found that 68.8% of the recruited mothers initiated early suckling of colostrum, but only 51.2% of the infants were exclusively breast-fed in the first week. Solid foods were introduced much earlier than at the recommended age of 4 to 6 months. Sixty percent of the mothers who participated in the study considered breast-feeding plus regular or irregular complementary feeding to be exclusive breast-feeding. (*Am J Public Health.* 1992;82:731–732)

Introduction

A high rate of breast-feeding is beneficial for the developing countries with regard to both economy and health. It is well known that colostrum improves the infant's chances of survival by providing both nutrition and immunization, specific and nonspecific, during the neonatal period.[1] Early suckling is therefore important to the infant and also stimulates the production of mature milk.[2]

Several studies[3–5] that distinguish between exclusive and partial breast-feeding have shown that a high percentage of infants in the developing countries are exclusively breast-fed for very short periods. The early introduction of water and food in environments where contamination is common not only exposes infants to concentrated numbers of microorganisms that may overwhelm the immunological protection provided by breast milk,[6] but also reduces stimulation of the breast, causing a reduction in the milk supply.[7]

Consequently, the study of infant feeding patterns during the neonatal period and the exclusive and partial breast-feeding periods is of crucial importance for an understanding of how to improve breast-feeding practices in the developing countries.

Methods

Mothers whose infants were born in 1989 and who attended the maternal and child health clinic in El-Kata, a rural village in Giza, Egypt, were asked if they would participate in a 1-year study of their infants' feeding patterns. The selected clinic and village are representative of rural Giza. The mothers of 250 infants were recruited.

At registration, the mothers were asked to provide information about themselves and their families, including the mother's level of education (no school, elementary school or less, or more than elementary school).

The mothers were interviewed for 1 year at weekly intervals from the time their infants were about 1 week old. At each interview, the mothers were asked about their infants' feeding. The mothers' definitions of "exclusive breastfeeding" were recorded in their own words.

The obtained data have been divided into 12-month equal intervals, with special interest in the first week of neonatal feeding. The Data Base III Plus program was used for data analysis.

Results

Of the 250 women enrolled in the study, 176 (70.4%) had had no schooling, 10 (4%) had an elementary-school education or less, and 64 (25.6%) had more than an elementary-school education.

Early suckling of colostrum was initiated by 68.8% of the mothers. We found that 61.4% of the uneducated mothers and 84.4% of the educated mothers (those with more than an elementary-school education) initiated early suckling of colostrum (Table 1). Causes of delayed suckling were found to be as follows: insufficient milk, 48%; traditional habits, 34.7%; tired mother, 6.2%; cesarean section, 4.1%; sick mother, 3.8%; weak baby, 3.1%.

The percentage of infants who were actually exclusively breast-fed in the first week was only 51.2%. The rest of them were given sugary fluids, honey and butter, or milk formula, or were fed by another lactating woman in the family. We found that 47.2% of the uneducated mothers and 70.3% of the educated mothers practiced exclusive breast-feeding in the first week.

Our results revealed that 128 infants (51.2%) were exclusively breast-fed for 1 week, 125 (50%) for 2 months, 106 (42.4%) for 3 months, 97 (38.8%) for 4 months, 86 (34.4%) for 5 months, 72 (28.8%) for 6 months, and 46 (18.4%) for 7 months. None of the infants was exclusively breast-fed after the age of 7 months; instead, they were given any available food, either regularly or irregularly.

The authors are with the Child Health Department, National Research Center, Giza, Egypt.

Requests for reprints should be sent to Iman A. Hakim, PhD, Child Health Department, National Research Center, Dokki, Giza, Egypt.

This paper was submitted to the Journal February 27, 1991, and accepted with revisions November 27, 1991.

TABLE 1—Percentages of Women Who Initiated Breast-feeding in the First Week

Start of Breast Suckling	Total (n = 250) No. (%)	Mother's Education: No School (n = 176) No. (%)	≤ Elementary (n = 10) No. (%)	> Elementary (n = 64) No. (%)
Colostrum				
Immediately after birth	60 (24.0)	52 (29.6)	0 —	8 (12.5)
After a few hours	61 (24.4)	26 (14.8)	5 (50.0)	30 (46.9)
Second day	44 (17.6)	28 (15.9)	0 —	16 (25.0)
Third day	7 (2.8)	2 (1.1)	5 (50.0)	0 —
Transitional				
Fourth day	22 (8.8)	22 (12.5)	0 —	0 —
Fifth Day	38 (15.2)	38 (21.6)	0 —	0 —
Sixth day	5 (2.0)	5 (2.8)	0 —	0 —
Seventh day	13 (5.2)	3 (1.7)	0 —	10 (15.6)

Only 40% of the mothers defined exclusive breast-feeding as giving breast milk only; 60% considered exclusive breastfeeding to include complementary feeding (giving breast-milk in conjunction with other foods). Also, 93.3% of the educated mothers described exclusive breast-feeding as giving breast milk with regular supplementary feeding (substituting 1 or more breast-feedings with other foods regularly) (Table 2).

Discussion

Mothers are encouraged to breast-feed their infants, but sooner or later, other foods and milk must be given to the infants.[8] In many cultures, colostrum is withheld or given sparingly.[9] In our study, the majority of the rural women (68.8%) initiated suckling of colostrum within the first 3 days; the remaining 31.2% started breast-feeding later. Only 51.2% of the infants were exclusively breastfed. The rest were given mixtures of sugar and water, honey and butter, tea and biscuit, formula, or were fed by another lactating woman in the family. Such foods were given in conjunction with breast milk for about 1 week or until mature milk became available. This practice was followed mainly because colostrum was said to be insufficient for the infant's energy needs and also because of the traditional belief that breast milk is not available for the neonate before 3 days, during which the infant may be fed by another lactating woman.

Maternal education, health behavior awareness, and increased capacity to change health belief systems are strongly correlated with initiation of breast-feeding.[10] In the present study, educated mothers were more likely to initiate early breast-feeding than were uneducated mothers. Similarly, infants of educated mothers were more likely than those of uneducated mothers to be exclusively breast-fed in the first week.

We found that 48.8% of the infants received not only mixtures of sugar and water, milk formula, and solids soon after birth and for about 1 week thereafter, but also complementary foods now and then until this became a regular practice. Solid foods were introduced much earlier than at the recommended age of 4 to 6 months[11] in 51.6% of the infants studied, often (in 3.2% of the infants) in the first month of life.

Mothers were frequently advised to introduce complementary foods on the basis of presumed milk insufficiency as soon as a problem, no matter how trivial, came up.[12] This pattern could be defined as regular complementary feeding or partial breast-feeding. Sixty percent of the mothers, however, described such a feeding pattern as exclusive breast-feeding.

It is crucial that we be aware of the very low prevalence of exclusive breast-feeding and the widespread practice of very early and irregular complementary feeding in the developing countries. Breast-feeding should be encouraged, but it is also important that irregular complementary feeding during the early weeks of life be strongly discouraged, unless such feedings are medically indicated. It seems likely that a change in this very early feeding pattern would contribute to a significant decline in the prevalence of neonatal systemic infection in the developing countries. □

TABLE 2—Mothers' Definitions of Exclusive Breastfeeding

Definitions	Total (n = 250) No. (%)	Mother's Education: No School (n = 176) No. (%)	≤ Elementary (n = 10) No. (%)	> Elementary (n = 64) No. (%)
Breast milk (BM)	100 (40)	96 (54.6)	0 —	4 (6.2)
BM & sugary fluid	35 (14)	28 (15.9)	7 (70)	0 —
BM & tea and biscuit	20 (8)	17 (9.7)	3 (30)	0 —
BM & available foods	25 (10)	25 (14.2)	0 —	0 —
BM & regular monthly supplements	70 (28)	10 (5.6)	0 —	60 (93.8)

References

1. Ogra PL, Dayton D. Immunology of breastmilk. New York: Raven Press; 1979.
2. Salariya EM, Easton PM, Cater JL. Duration of breastfeeding after early initiation and frequent feeding. *Lancet.* 1978;11:1141.
3. Forman MR. Review of research on the factors associated with choice and duration of infant feeding in less-developed countries. Task force on infant feeding practices. *Pediatrics.* 1984;74(4Pt2):667–694.
4. Hijazi SS, Abulaban A, Waterlow JC. The duration for which exclusive breastfeeding is adequate. *Acta Paediatr Scand.* 1989;78:23–28.
5. Ashley D, Fox K. An assessment of socio-economic factors associated with the pattern of breastfeeding in Jamaica. *Int Pediatr Assoc Bull.* 1983;5:30–37.
6. Hanson LA, Hofvander Y, Lindquist B, Zetterström R. Breastfeeding as a protection against gastroenteritis and other infections. *Acta Paediatr Scand.* 1985;74:641–642.
7. Egli GE, Newton M. The influence of number of breastfeeding on milk production. *Pediatrics.* 1961;27:314–317.
8. Eaton-Evans J, Dugdale AE. Effects of feeding and social factors on diarrhea and vomiting in infants. *Arch Dis Child.* 1987;6:445–448.
9. Jellife DB, Jellife EEP. Human milk in the modern world. Oxford, England: Oxford University Press; 1978.
10. Becerra JE, Smith JC. Breastfeeding patterns in Puerto Rico. *Am J Public Health.* 1990;80:694–697.
11. Underwood BA, Hofvander Y. Appropriate timing in complementary feeding of the breastfed infant. A review. *Acta Paediatr Scand.* 1982;71(Suppl. 294).
12. Koctürk T. Events leading to the decision to introduce complementary feeding to the breast among a group of mothers in Istanbul. *Scand J Prim Health Care.* 1986; 4:231–237.

Access to Cancer Screening Services for Women

Bradford Kirkman-Liff, DrPH, and Jennie Jacobs Kronenfeld, PhD

ABSTRACT

A major effort in preventive care for women has emphasized the obtaining of Pap smears and mammograms. This paper uses survey data from one state to examine issues of access to Pap smears and mammograms. Poor women receiving health care through a managed-care Medicaid program received these services at the same rate as women with other types of health insurance, while the uninsured were less likely to have had either type of service. (*Am J Public Health.* 1992;82:733–735)

Introduction

One preventive care emphasis for women is obtaining Pap smears and mammograms to detect cancer at an earlier, more curable stage. Recent studies have particularly linked poverty to cancer risk.[1–3] This paper uses survey data to examine issues of access to Pap smears and mammograms. Of particular interest is whether poor women receiving health care through the Arizona Health Care Cost Containment System (AHCCCS), a unique, managed-care Medicaid program, receive these services at the same rate as women with other health insurance. Earlier studies reported that poverty, age, no health insurance, and no recent visit to an obstetrician/gynecologist related to lack of Pap tests and that poverty, low education, and age related to lack of mammograms.[3–9] This analysis compares AHCCCS women with uninsured women and women with other types of insurance. Unique to this analysis is the inclusion of use of each service (Pap smear, mammogram) in the model for the other service. Previous studies have generally focused on only one aspect of screening services.

Until 1981, Arizona had no Medicaid program. AHCCCS, a managed-care/health maintenance organization alternative to traditional fee-for-service Medicaid, was designed to provide cost-effective care that avoids excessive costs from overuse. The program operates through a network of capitated, managed-care plans with primary-care physicians as gatekeepers who coordinate all needed services. Plans competitively bid and negotiate for contracts and are at financial risk for overuse. Eligible recipients must obtain all care from their selected plan.

Methods

Survey data were collected in 1989. About 3100 randomly selected adults were interviewed about themselves, and 1100 about a randomly selected child. Interviews were conducted in English or Spanish. Further details on this survey have been published elsewhere.[10] Four measures of access to cancer screening services were examined. In addition to a dichotomous variable of whether a woman ever had a test, women who had been screened by each test were asked when it occurred. Women who had received each service were split into those receiving it within 2 years prior to the survey and those 3 years or more before the survey. We also examined whether women who had a Pap smear were more likely to have a mammogram, and vice versa. Mammogram analyses included only women over 40, since mammogram services are not recommended at regular intervals before age 40.[11]

In addition to examining the impact of insurance status and the AHCCCS program, other variables were hypothesized to affect use of preventive services. The different cultural values of Latino women may lead to different use patterns. (Numbers are too small to allow analyses of African-American and Native American women.) Latino women interviewed in Spanish were expected to have differences from those interviewed in English, with language of interview serving as an acculturation proxy. Educational attainment would affect receptivity to cancer screening messages, with less-educated women reporting lower use. Women with no usual health care source would have lower use of cancer screening services than women with a usual source. Women who needed health care but were unable to obtain it were expected to have lower Pap smear/mammogram use. Number of visits with health providers was expected to increase the likelihood of receiving cancer screening services, possibly nonlinearly. Lastly, age would have a varying effect, depending on the specific service.

The authors are with the School of Health Administration and Policy, College of Business, Arizona State University, Tempe.

Requests for reprints should be sent to Bradford Kirkman-Liff, DrPH, School of Health Administration and Policy, College of Business, Arizona State University, Tempe, AZ 85287-4506.

This paper was submitted to the Journal June 25, 1991, and accepted with revisions December 9, 1991.

TABLE 1—Occurrence of Pap Smear, Arizona Anglo and Latino Women, 1989

Measure	Pap Smear within Last 2 Years, %	Pap Smear 3 or More Years Ago, %	Never Had Pap Smear, %	χ^2	P
Insurance status				16.45	.002
No insurance	44.3	42.2	13.5		
AHCCCS	72.9	20.4	6.7		
Other insurance	69.9	21.3	8.8		
Education				13.92	.008
Less than high school	62.0	23.3	14.7		
High school	67.6	23.4	9.1		
More than high school	74.9	20.8	4.4		
Usual source of care				27.45	<.001
No	51.9	33.4	14.7		
Yes	72.9	19.7	7.4		
Need but not receive medical care				7.52	.023
Yes	53.4	33.3	13.3		
No	69.6	21.8	8.6		
Ethnicity and language of interview				2.48	.659
Anglo	67.8	23.7	8.5		
Latino and English interview	70.5	20.0	9.5		
Latino and Spanish interview	78.1	11.0	10.9		
All respondents	68.1	22.8	9.0		

Note. AHCCCS = Arizona Health Care Cost Containment System.

TABLE 2—Occurrence of Mammogram, Arizona Anglo and Latino Women, 1989

Measure	Mammogram within Last 2 Years, %	Mammogram 3 or More Years Ago, %	Never Had Mammogram, %	χ^2	P
Insurance status				44.51	<.001
No insurance	9.0	7.2	83.2		
AHCCCS	53.6	3.9	42.5		
Other insurance	49.6	11.5	38.9		
Education				4.76	.313
Less than high school	42.4	10.0	47.6		
High school	46.8	10.2	42.9		
More than high school	50.7	12.5	36.9		
Usual source of care				24.51	<.001
No	30.4	14.1	55.6		
Yes	51.8	9.8	38.4		
Need but not receive medical care				5.74	.057
Yes	36.1	18.3	45.5		
No	47.9	10.1	42.0		
Ethnicity and language of interview				7.80	.099
Anglo	47.6	11.6	40.8		
Latino and English interview	45.2	7.1	47.7		
Latino and Spanish interview	34.1	1.5	64.5		
All respondents	46.9	10.8	42.3		

Note. AHCCCS = Arizona Health Care Cost Containment Service.

Results

Overall, 68% of women had a Pap smear in the 2 years prior to the survey, 23% had a Pap smear 3 years or more prior to the study, and 9% never had a Pap smear. Exploratory cross tabulations showed statistically significant differences by insurance status, education, usual source of care, and needing but not receiving medical care (Table 1). Overall, 47% of women had a mammogram in the 2 years prior to the survey, 11% had a mammogram 3 years or more prior to the study, and 42% never had a mammogram. Only insurance status and usual source of care were significantly related to use of mammograms (Table 2).

Logistic regression was used to model the dichotomous measures. The results were transformed into odds ratios and 95% confidence levels. Variables without significant odds ratios are not presented, except for insurance. Women who had a mammogram were far more likely to have had a Pap smear than women without a mammogram, and women who had a Pap smear were far more likely to have had a mammogram than women without a Pap smear. Uninsured women and women with no usual source of care were far less likely to have had either service than women with conventional insurance, in AHCCCS, or with a usual source of care. Latino women were less likely to report a Pap smear, with no difference in mammograms (Table 3).

Women who had a mammogram were far more likely to have had a Pap smear within 2 years of the survey than women without a mammogram, and vice versa. Uninsured women and women with no usual source of care were far less likely to have had either service within 2 years of the survey than women with other insurance, AHCCCS, or a usual source of care (Table 4).

Discussion and Conclusion

Several important conclusions emerge from these analyses. The insignificant AHCCCS term indicates that a managed-care Medicaid-type model led to more appropriate use of preventive health services in cancer screening. AHCCCS women had no difference in access to these two preventive health services as compared with women with other insurance. Uninsured women and those with no usual source of care were less likely to have ever had the services or to have had them within the previous 2 years. In addition, older women and Latino women were less likely to have ever received a Pap smear. A recent editorial about race, poverty, and cancer concluded that poor Americans constitute a high-risk group for cancer.[12] This paper demonstrates that managed-care Medicaid programs can

TABLE 3—Odds Ratios on Service Use

Model/Variable	Odds Ratio	95% Confidence Interval
Have had a Pap smear		
AHCCCS	0.83	0.41–1.70
Uninsured	0.44	0.32–0.60
Latino	0.56	0.44–0.72
No usual source of care	0.61	0.49–0.76
60–70 years old	0.36	0.27–0.49
70 and older	0.19	0.14–0.25
Have had a mammogram	3.73	2.93–4.75
Have had a mammogram		
AHCCCS	0.95	0.60–1.52
Uninsured	0.29	0.23–0.38
No usual source of care	0.55	0.46–0.66
No. of visits (relative to a woman with no visits)		
1	1.03	1.02–1.04
2	1.07	1.04–1.10
4	1.15	1.09–1.20
7	1.26	1.16–1.38
10	1.39	1.22–1.58
20	1.79	1.34–2.40
70 and older	0.69	0.59–0.81
Have had a Pap smear	3.64	2.86–4.63

Note. AHCCCS = Arizona Health Care Cost Containment Service.

TABLE 4—Odds Ratios on Recency of Service Use

Model/Variable	Odds Ratio	95% Confidence Interval
Had Pap smear within last 2 years		
AHCCCS	0.75	0.44–1.29
Uninsured	0.54	0.41–0.70
No usual source of care	0.54	0.44–0.65
Needed but did not receive care	0.51	0.40–0.65
60–70 years old	0.49	0.40–0.61
70 and older	0.35	0.29–0.44
Have had a mammogram	3.12	2.63–3.70
Had mammogram within last 2 years		
AHCCCS	1.42	0.63–3.17
Uninsured	0.36	0.23–0.56
No usual source of care	0.38	0.29–0.51
Have had a Pap smear	2.44	1.59–3.75

Note. AHCCCS = Arizona Health Care Cost Containment Service.

help direct resources to people at higher risk. Some studies have emphasized the need for motivational programs to encourage usage.[4] In contrast, this study provides evidence that, without special motivational programs, if poor women receive health care services within a managed-care model with a preventive approach they have similar rates of use as the general population of insured women. □

Acknowledgments

This research was supported in part by the Flinn Foundation, Phoenix, Ariz.

All opinions and views expressed are solely those of the authors and do not necessarily represent those of the Flinn Foundation.

References

1. Baquet C, Horn J, Gibbs T, Greenwald P. Socioeconomic factors and cancer incidence among Blacks and Whites. *JNCI.* 1991;83:551–559.
2. Howard J. Avoidable mortality from cervical cancer: exploring the concept. *Soc Sci Med.* 1987;24:507–514.
3. Baquet C, Ringen K. Health policy: gaps in access, delivery, and utilization of the Pap smear in the United States. *Milbank Q.* 1987;65:322–347.
4. Mamon J, Shediac M, Crosby C, Sanders B, Matanoski G, Celentano D. Inner city women at risk for cervical cancer: behavioral and utilization factors related to inadequate screening. *Prev Med.* 1990;19:363–376.
5. Richardson J, Marks G, Solis J, Collins J, Birba L, Hisserich J. Frequency and adequacy of breast cancer screening among elderly Hispanic women. *Prev Med.* 1987;16:761–774.
6. Centers for Disease Control. Use of mammography for breast cancer screening—Rhode Island, 1987. *MMWR.* 1988;37:347–360.
7. Anda R, Sienko D, Remington P, Gentry E, Marks J. Screening mammography for women 50 years of age and older: practices and trends, 1987. *Am J Prev Med.* 1990;6:123–129.
8. Mandelblatt J, Gopaul I, Wistreich M. Gynecological care of elderly women. *JAMA.* 1986;256:367–371.
9. Zapka JG, Stoddard A, Maul L, Costanza ME. Internal adherence to mammography screening guidelines. *Med Care.* 1991;29:697–707.
10. Kirkman-Liff B, Mondragon D. Language of interview: relevance for research of Southwest Hispanics. *Am J Public Health.* 1991;81:1399–1404.
11. Eddy D, Hasselblad V, McGivney W, Hendee W. The value of mammography screening in women under age 50 years. *JAMA.* 1988;259:1512–1519.
12. Freeman H. Race, poverty and cancer. *JNCI.* 1991;83:526–527.

Moderate Physical Activity and Bone Density among Perimenopausal Women

Jun Zhang, MB, Paul J. Feldblum, MSPH, and Judith A. Fortney, PhD

ABSTRACT

We examined the association between physical activity and bone mineral density (BMD) among 352 perimenopausal women. A personal activity computer was used to estimate the daily energy expenditure on physical activity. BMD was measured by photon absorptiometry at the lumbar spine, midradius, and distal radius. Multiple linear regression, controlling for other factors, indicated that physical activity was positively and significantly associated with BMD at all sites. Moderate physical activity has beneficial effects on BMD among perimenopausal women. (*Am J Public Health.* 1992;82:736–738)

Introduction

It has been suggested that exercise increases bone density[1] and prevents or decelerates age- and menopause-related bone loss.[2] Studies comparing highly trained athletes and sedentary individuals concluded that exercise increases bone density.[3–6] However, conflicting findings were reported from both cross-sectional and prospective studies with nonathlete subjects.[7,8] Block et al. suggested that inaccurate or subjective measurement of physical activity by questionnaires may, in part, account for the conflicting findings.[9] Small study sizes, resulting in low statistical power, may be another important factor.

The present study used personal activity computers to measure physical activity objectively. These devices are considered by some researchers to provide superior direct measurements of physical activity.[10,11] In addition to its large size, our study also adjusted for other determinants of bone density.

Materials and Methods

A group of 352 White women 40 to 54 years of age were recruited in central North Carolina by means of newspaper advertisement and flyers in workplaces, stores, and doctors' offices. Volunteers were required to have at least one ovary and no physical impairment inhibiting daily activities. Women who were pregnant, had smoked within the past 5 years, or were using estrogen replacement therapy were ineligible. Also excluded were women with a history of acromegaly, parathyroid disease, insulin-dependent diabetes mellitus, inflammatory bowel disease, liver disease (except hepatitis), or renal diseases (except pyelonephritis).

Participants completed a self-administered questionnaire that covered health history, reproductive and contraceptive histories, past and current exercise, and past and current diet. A simple life calendar was provided to aid recall. Menopausal status was defined on the basis of the combination of answers to items on menstrual changes, recent menses, and last menses. A woman was designated postmenopausal if she (1) reported menstrual changes, (2) reported no menstrual period in the previous 6 months, and (3) described herself as being in menopause. Others were designated premenopausal.

A series of questions concerning the respondents' exercise in high school were asked. Adolescent exercise was coded as follows: high level—played sports on a school team; moderate level—exercised regularly; and low level—did not exercise regularly.

A current food intake frequency chart was included in the questionnaire, from which dietary calcium was calculated.[12] We also queried the women about nondietary calcium in the form of antacids and dietary supplements. Based on the criterion of the recommended dietary allowance, subjects were dichotomized as above or below recommended total daily calcium intake.

Current physical activity was measured by a personal activity computer (Caltrac, Hemokinetics Inc, Madison, Wis), which is a portable accelerometer estimating daily energy expenditure in kilocalories.[11] The subjects were instructed to wear the device for 3 continuous days (all day long), including at least 1 weekend day and 1 weekday. Daily energy expenditure on physical activity was derived from total daily energy expenditure minus basal metabolic rate (BMR): BMR = [(331 × weight) + (351 × height) − (352 × age) + 49 854]/100 000 (kcal/min), where weight is measured in pounds, height in inches, and age in years.

The authors are with the Division of Reproductive Epidemiology and Sexually Transmitted Diseases, Family Health International, Research Triangle Park Branch, Durham, NC.

Requests for reprints should be sent to Paul J. Feldblum, MSPH, Division of Reproductive Epidemiology and Sexually Transmitted Diseases, Family Health International, PO Box 13950, Research Triangle Park Branch, Durham, NC 27709.

This paper was submitted to the Journal May 17, 1991, and accepted with revisions November 12, 1991.

TABLE 1—Bone Mineral Density (mg/cm²) by Levels of Explanatory Variables

Feature	n	Spine L_2–L_4		Distal radius		Midradius	
		Mean	SD	Mean	SD	Mean	SD
Age							
40–44	151	1131	137	423	54	753	56
45–49	115	1126	143	415	57	760	52
50–54	86	1092	166	415	66	745	56
Parity							
0	68	1090	126	413	64	754	49
1	52	1127	136	419	60	747	65
2	124	1133	151	420	46	753	51
3+	108	1121	159	420	64	757	58
Breast-feeding, mo							
None	148	1090	145	412	61	750	52
1–12	111	1149	140	422	56	752	60
13+	93	1130*	153	424	54	761	51
Oral contraceptive use, y							
None	92	1114	143	417	63	751	59
1–2	87	1130	152	424	59	759	56
3–5	79	1110	137	419	57	754	58
6+	94	1125	156	414	52	750	48
Exercise in high school							
Low	129	1113	139	412	51	751	51
Moderate	126	1119	142	420	58	751	60
High	97	1131	164	424	65	759	53
Calcium intake[a]							
Above RDA	198	1114	144	419	59	754	55
Below RDA	154	1126	151	418	56	753	55
Menopausal status							
Premenopausal	264	1133	139	421	58	755	54
Postmenopausal	88	1082*	165	411	58	748	57
Physical activity, kcal/day[b]							
Low (<316)	116	1102	139	412	56	744	53
Medium (317–488)	117	1094	141	408	54	756	52
High (>488)	119	1163**	152	435**	61	760	59

Note. L_2–L_4 = second to fourth lumbar vertebrae. RDA = recommended dietary allowance.
[a]Calcium intake was calculated on the basis of US Department of Agriculture and RDA criteria.
[b]Tertiles of the personal activity computer mean reading, after removing the effects of age and body mass.
*$P < .01$.
**$P < .001$.

Bone mineral density (BMD) was measured by dual photon absorptiometry at the second to fourth lumbar vertebrae, and single photon absorptiometry measured BMD at the midradius and the distal radius (5-mm site).

Results

The spine BMD slightly decreased with age, and postmenopausal women had substantially lower lumbar BMD than did premenopausal women (Table 1). Women who had ever breast-fed had significantly higher lumbar BMD than did nonbreast-feeding women. There was a modest dose–response relationship between exercise in high school and BMD at all sites that did not reach statistical significance. Women with more physical activity as measured by the personal activity computer had significantly higher bone density at the spine and distal radius. This association also prevailed at the midradius but was not statistically significant.

Multiple linear regression, controlling for age, body mass, breast-feeding, oral contraceptive use, exercise in high school, and menopausal status, indicated that current physical activity is directly and significantly associated with bone density at the spine, distal, and midradius sites (Table 2). Based on these coefficients, each additional 100 kcal/day energy expenditure in physical activity adds an estimated 12 mg/cm², 3.8 mg/cm², and 3.5 mg/cm² higher bone density at the spine, distal, and midradius sites, respectively.

Discussion

It has been established that highly trained athletes have significantly higher BMD than do sedentary persons.[3–6] Although research on this highly selected group suggests that intensive exercise increases BMD, the finding has lesser practical and public health value for perimenopausal women, for whom the issue of moderate physical activity may be of greater interest. The present study shows that perimenopausal women with more physical activity have significantly higher BMD when other determinants of BMD are taken into account.

Some limitations of our study should be noted. First, this was a cross-sectional study in which limited information on past exercise, body weight, and diet was collected. Second, misclassification may exist for some variables such as exercise in high school. Behavior change after adolescence may further lessen the value of this indicator. Third, although we improved the measurement of physical activity by using personal activity computers, it can be argued that 3 days of physical activity are not truly representative of the

TABLE 2—Multiple Linear Regression Analysis of the Association between Physical Activity and Bone Mineral Density

Site	Beta Coefficient[a]	SE	95% Confidence Interval	*P* Value
Spine L_2–L_4	0.1203	0.0386	0.0443–0.1963	.0021
Distal radius	0.0375	0.0161	0.0060–0.0691	.0207
Midradius	0.0350	0.0156	0.0040–0.0656	.0261

Note. L_2–L_4 = second to fourth lumbar vertebrae.
[a]Multiple linear regression analysis adjusted for age, Quetelet index, months of breast-feeding, months of oral contraceptive use, exercise in high school (ordinal), and menopausal status (dichotomous). Quetelet index = weight (lb) × 100/height (in)^2.

women's life-styles. Some women might have increased their exercise in the 3 measured days because of the study. Furthermore, since the device could only measure acceleration and deceleration in a vertical direction, some types of exercise, such as bicycling, may accumulate too few kilocalories. If this were true, however, the bias would be in the direction of the null because of increased variance. Thus, our significant findings suggest either that there is no strong bias in the physical activity results in our study or that the relationship between activity and BMD was underestimated.

Nonetheless, the advantages of the present study are noteworthy. The study size was substantially larger than most previous studies of exercise and BMD. Use of personal activity computers provided both objectivity and precision. In addition, the present study measured and adjusted for factors that were often overlooked in previous studies, such as breast-feeding, oral contraceptive use, and calcium intake. These advantages, along with the significant findings of this study and their biological plausibility and consistency with previous studies, lead us to conclude that moderate physical activity is associated with higher bone density among perimenopausal women. □

Acknowledgments

Partial support for this work was provided by Family Health International (FHI) with funds from the US Agency for International Development (USAID) and the National Institute for Child Health and Human Development (NICHD), although the views expressed in this article do not necessarily reflect those of USAID or NICHD. FHI is a not-for-profit organization that conducts research and provides technical assistance in reproductive health, family planning, and sexually transmitted diseases.

References

1. Dalsky GP, Stocke KS, Ehsani AA, Slatopolsky E, Lee WC, Birge SJ. Weight-bearing exercise training and lumbar bone mineral content in postmenopausal women. *Ann Intern Med.* 1988;108:824–828.
2. Chow R, Harrison JE, Notarius C. Effect of two randomised exercise programmes on bone mass of healthy postmenopausal women. *Br Med J.* 1987;295:1441–1444.
3. Jacobson PC, Beaver W, Grubb SA, et al. Bone density in women: college athletes and older athletic women. *J Orthop Res.* 1984;2:328–332.
4. Wolman RL, Clark P, McNally E, Harries M, Reeve J. Menstrual state and exercise as determinants of spinal trabecular bone density in female athletes. *Br Med J.* 1990;301:516–518.
5. Lane NE, Bloch D, Jones HH, et al. Long-distance running, bone density, and osteoarthritis. *JAMA.* 1986;255:1147–1151.
6. Block JE, Genant HK, Black DM. Greater vertebral bone mineral mass in exercising young men. *West J Med.* 1986;145:39–42.
7. Simkin A, Ayalon J, Leichter I. Increased trabecular bone density due to bone-loading exercises in postmenopausal osteoporotic women. *Calcif Tissue Int.* 1986;39:8.
8. Aloia JF, Cohn SH, Ostuni JA, et al. Prevention of involutional bone loss by exercise. *Ann Intern Med.* 1978;89:356–358.
9. Block JE, Smith R, Black D, Genant HK. Does exercise prevent osteoporosis? *JAMA.* 1987;257:3115–3117.
10. Washburn RA, Montoye HJ. The assessment of physical activity by questionnaire. *Am J Epidemiol.* 1986;123:563–576.
11. Montoye HJ, Washburn R, Servais S, Ertl A, Webster JG, Nagle FJ. Estimation of energy expenditure by a portable accelerometer. *Med Sci Sports Exerc.* 1983;15:403–407.
12. US Department of Agriculture, Agricultural Research Service. *Composition of Foods, Dairy and Egg Products: Raw, Processed, Prepared.* Washington, DC: US Government Printing Office; 1976. Agriculture Handbook No. 8-1.

Medical Expenditures for Disability and Disabling Comorbidity

Dorothy P. Rice, ScD (Hon.), and Mitchell P. LaPlante, PhD

ABSTRACT

Disability and disabling comorbidity place a disproportionately large burden on the health care system. National Medical Care Utilization and Expenditure Survey data show that medical care expenditures for noninstitutionalized persons amounted to $154 billion ($691 per capita) in 1980. The medical expenditure per capita for people reporting two or more disabling chronic conditions ($2456) was 5 times the amount incurred by those with no limiting conditions ($486) and more than 1.5 times the amount incurred by those with one limiting condition ($1620). (*Am J Public Health.* 1992;82:739–741)

Introduction

Comorbidity, defined as the coexistence of multiple diseases and chronic conditions, has long been recognized as an important contributing factor in death.[1,2] The negative impact of comorbidity on health and functional ability has not been well documented until recently. A body of research is now emerging that focuses on the prevalence of comorbidity and its health consequences in representative samples of the population.[3–8] Data from the National Health Interview Survey show that disability and disabling comorbidity have increased recently. From 1970 to 1980, larger proportions of people were disabled by chronic conditions that limited their activities, and such persons reported more activity-limiting conditions per capita.[9]

Disability and disabling comorbidity place a large burden on the health care system, but the magnitude of the medical care costs involved has not been measured. This paper estimates the total and per capita medical expenditures in 1980 by number of activity-limiting chronic conditions and by age and type of expenditure.

Methods

The data on medical expenditures described here are taken from the 1980 National Medical Care Utilization and Expenditure Survey (NMCUES) conducted by the National Center for Health Statistics. A national probability sample of about 6000 households was augmented by samples drawn from the Medicaid eligibility rolls of California, Michigan, New York, and Texas (1000 households in each state). Data collection consisted of initial interviews during February through April 1980 and four rounds of follow-up interviews at approximately 3-month intervals.[10]

Information was collected on a wide range of topics related to health status, access to and use of medical services, associated charges, sources of payment, and health insurance coverage. This paper presents data obtained from the NMCUES public use data files, which contain only respondent data from the National Household Survey. Total charges were obtained for medical services and supplies in the survey. Missing information for critical data was imputed. Total charges for medical visits, hospital stays, and prescribed medicines and other medical expenses were imputed for 25.9%, 36.3%, and 19.4% of the events, respectively.[11] The methods used to impute for missing items were diverse and tailored to the measure requiring information.

Data are reported here on medical expenditures for persons with one and two or more limiting chronic conditions as assessed in Round 1. The number of persons reporting more than two conditions was not large enough to permit reliable estimates. A chronic condition is defined as having an actual or expected duration of more than 3 months. To be classified as limited, a person must report one or more chronic conditions as a cause of limitation of activity defined in terms of the major activities usually associated with a person's age group: (1) ordinary play for children under 5 years of age, (2) attending school for those 5 through 17 years of age, and (3) working or keeping house for persons aged 18 and over. Questions on functional limitations were asked only in the last interview (Round 5) and were not used in this analysis, although we recognize the utility of such measures, especially for older persons.[12]

Results

Of the 223 million noninstitutionalized persons in 1980, 36 million, or 16.2%, reported limitation of activity due to chronic conditions. About 85% (30.5 million persons) of those reporting activity limitations had one condition and the remaining 15% (5.5 million persons) had two or more conditions (Table 1).

The authors are with the Institute for Health & Aging, School of Nursing, University of California, San Francisco.

Requests for reprints should be sent to Dorothy P. Rice, Institute for Health & Aging, School of Nursing, N 631 Y, University of California, San Francisco, CA 94143-0612.

This paper was submitted to the Journal March 18, 1991, and accepted with revisions August 22, 1991.

TABLE 1—Number of Persons and Medical Expenditures by Age and Number of Limiting Chronic Conditions, 1980

No. of Conditions and Age	No. of Persons (thousands)	Percentage Distribution	Medical Expenditures		
			Total (millions)	Percentage Distribution	Per Capita
Total	222 818	100.0	$153 863	100.0	$ 691
<19	67 972	100.0	21 705	100.0	319
19–44	86 092	100.0	50 911	100.0	591
45–64	43 578	100.0	39 713	100.0	911
≥65	25 176	100.0	41 533	100.0	1650
No conditions	186 797	83.8	90 856	59.0	486
<19	64 841	95.4	19 468	89.7	300
19–44	75 785	88.0	36 543	71.8	482
45–64	31 557	72.4	18 542	46.7	588
≥65	14 614	58.0	16 303	39.3	1116
One condition	30 468	13.7	49 369	32.1	1620
<19	2999	4.4	2202	10.1	734
19–44	9424	10.9	12 925	25.4	1371
45–64	9735	22.3	16 427	41.4	1687
≥65	8309	33.0	17 815	42.9	2144
Two or more conditions	5554	2.5	13 638	8.9	2456
<19	132	0.2	35*	0.2	267*
19–44	882	1.0	1444	2.8	1637
45–64	2286	5.2	4744	11.9	2075
≥65	2254	9.0	7415	17.9	3290

Note. Numbers and percentages may not add up to totals as a result of rounding.
*Relative standard error exceeds 30%.
Source: 1980 National Medical Care Utilization and Expenditure Survey. Authors' original tabulations from public use tapes.

Total expenditures for medical care for all noninstitutionalized persons amounted to $153.9 billion, or $691 per person, in 1980. The 187 million people with no activity limitations incurred $90.9 billion, or $486 per capita. Thus, 84% of those who reported no limitations incurred 59% of the total medical expenditures. By contrast, 14% of the people with one limiting condition incurred 32% of the total expenditures, while the less than 3% of those reporting two or more conditions incurred almost 9% of the total expenditures. On a per capita basis, expenditures for persons with one condition amounted to $1620, over three times the amount spent per capita for persons with no activity limitations. For those reporting two or more conditions, the medical expenditure per capita was $2456, 5 times the amount incurred by those with no limiting conditions ($486) and more than 1.5 times the amount incurred by those with one limiting condition ($1620).

Costs by Age

Total per capita expenditures increased fivefold with age, rising from $319 for those under age 19 to $1650 for the elderly aged 65 and over. More than 11% of the population were aged 65 and over in 1980, and this group incurred 27% of medical expenditures. Persons under 19 years of age constituted 31% of the population but accounted for only 14% of the charges.

This pattern of higher per capita charges for older persons also applies to those limited in activity because of one or more chronic conditions. For those reporting one condition, per capita charges rose from $734 for the youngest age group (under 19 years) to $2144 for the elderly. For those reporting two or more conditions, per capita expenditures rose from $1637 for those aged 19 through 44 to $3290 for those aged 65 and over. (There were too few cases under age 19 who reported two or more chronic conditions to produce reliable per capita charges.)

Comorbidity is often perceived as a problem associated with the aged population. Comorbidity rates among the noninstitutionalized population under age 65 were lower than for the elderly (1.7% and 9.0%, respectively, reported limitation of activity due to multiple chronic conditions). Nevertheless, the economic burden of comorbidity for the under-65 age group is substantial. Of the $37.8 billion in total medical expenditures generated by persons under age 65 reporting one or more limiting conditions, $6.2 billion, or 16.5%, was spent in behalf of persons reporting two or more conditions. For the elderly, $7.4 billion of the $25.2 billion, or 29.4%, involved those with two or more limiting conditions. Thus, those under age 65 generate 45.6% of the $13.6 billion of total medical costs for disabling comorbidity.

Type of Expenditure

Of the total medical expenditures amounting to $153.9 billion in 1980, almost one half (47%) involved hospital inpatient care and about 10% involved physician inpatient services. Outpatient services accounted for more than 20% of the total expenditures: $11.3 billion (7%) for hospital outpatient services and $20.6 billion (13%) for physician services provided outside the hospital. Charges for dental and other professional services amounted to $23.3 billion, 15% of total medical expenditures; charges for prescription drugs amounted to $7.8 billion (5%); and $3.7 billion (2%) was spent for other medical expenses (Table 2).

For each type of expenditure except for dental services, per capita charges for those with no limiting conditions were significantly lower than for persons reporting one or more limiting condition. For example, per capita charges for inpatient care for persons with one condition were five times the amount spent for those without limitations. For persons with two or more conditions, per capita hospital inpatient charges were eight times the amount spent for those without limitations. More than three fifths of the costs of comorbidity were due to inpatient hospitalization. It is clear that the existence of comorbid conditions results in longer hospital stays and higher overall costs.

Conclusions

This analysis clearly shows that medical care costs for disability and disabling comorbidity impose a large burden on society. The estimates presented here quantify only one portion of the economic and social costs. In addition to medical resources used, disability also results in reduced or lost productivity (estimated at $68.4 billion in 1980[13]), premature death, and pain and suffering of persons with disabilities, their families, and their friends. Considerable savings to society can be gained if the incidence and prevalence of disability and disabling comorbidity can be prevented or reduced. Comorbidity among people with disabilities can be prevented given appropriate access to care and health promotion.[14] The 1990 Americans with Disabilities Act was designed to improve access to jobs, transportation, housing, entertainment,

TABLE 2—Medical Expenditures by Number of Limiting Conditions and Type of Expenditure, 1980

Type of Expenditure	All Persons		Not Limited		One Condition		Two or More Conditions	
	Total (millions)	Per Capita	Total (millions)	Per Capita	Total (millions)	Per Capita	Total (millions)	Per Capita
Total	$153 863	$691	$90 856	$486	$49 369	$1620	$13 638	$2456
Hospital inpatient	71 958	323	35 736	191	27 644	907	8578	1545
Hospital outpatient	11 273	51	7237	39	3135	103	902[a]	162[a]
Physician inpatient	15 183	68	8444	45	5348	176	1391	250
Physician outpatient	20 605	92	14 115	76	5391	177	1099[a]	198[a]
Dental	17 691	79	15 046	81	2304	76	341[a]	61[a]
Other professional	5627	25	3730	20	1670	55	228[a]	41[a]
Prescribed medicines	7831	35	4120	22	2887	95	824[a]	148[a]
Other medical	3714	17	2436	13	999	33	279	50

[a]Relative standard error exceeds 30%.
Source: 1980 National Medical Care Utilization and Expenditure Survey. Authors' original tabulations from public use tapes.

and other activities and to provide an important social, economic, and legal impetus to improve the lives of persons with disabilities. The recent Institute of Medicine report, *Disability in America: Toward a National Agenda for Prevention,* emphasized the need for more research in preventing disability and disabling comorbidity.[15] □

Acknowledgments

Support for this research was provided by the National Institute for Disability and Rehabilitation Research, US Department of Education, Washington, DC (cooperative agreement no. G0087C 2014-89).

This paper was originally presented at the 116th Annual Meeting of the American Public Health Association, November 1988, Boston, Mass.

References

1. Guralnick L. Some problems in the use of multiple causes of death. *J Chronic Dis.* 1966;19:979–990.
2. Israel RA, Rosenberg HM, Curtin LR. Analytic potential for multiple cause-of-death data. *Am J Epidemiol.* 1986; 124:161–179.
3. Seeman TE, Guralnik JM, Kaplan GA, Knudsen L, Cohen R. The health consequences of multiple morbidity in the elderly. *J Aging Health.* 1989;1:50–66.
4. Guralnik JM, LaCroix AZ, Everett DF, Kovar MG. *Aging in the Eighties: The Prevalence of Comorbidity and Its Association with Disability.* Hyattsville, Md: National Center for Health Statistics; 1989. (Advance data from Vital and Health Statistics; no. 170).
5. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chronic Dis.* 1987;40:373–383.
6. Verbrugge LM, Lepkowski JM, Imanaka Y. Comorbidity and its impact on disability. *Milbank Q.* 1989;67:450–484.
7. Regier DA, Farmer ME, Rae DS, et al. Comorbidity of mental disorders with alcohol and other drug abuse. *JAMA.* 1990;264:2511–2518.
8. LaPlante MP. *Data on Disability from the National Health Interview Survey, 1983–85.* An InfoUse Report. Washington, DC: US National Institute on Disability and Rehabilitation Research; 1988.
9. Rice DP, LaPlante MP. Chronic illness, disability, and increasing longevity. In: Sullivan S, Lewin ME, eds. *The Economics and Ethics of Long-Term Care and Disability.* Washington, DC: American Enterprise Institute for Public Policy Research; 1988:9–55.
10. Bonham GS. *Procedures and Questionnaires of the National Medical Care Utilization and Expenditure Survey.* Washington, DC: US Government Printing Office, 1983. Series A, Methodological Report No. 1. DHHS publication no. 83–20001, National Center for Health Statistics, Public Health Service.
11. Parsons PE, Lichtenstein R, Berki SE, et al. *Costs of Illness, United States, 1980. National Medical Care Utilization and Expenditure Survey.* Washington, DC: US Government Printing Office; 1986. Series C, Analytical Report No. 3. DHHS publication no. 86–20403, National Center for Health Statistics, Public Health Service.
12. Branch GB, Richmond JB, Rogers DE, Wilson RW, Adams ME. Health indicators for an aging population. In: *Health of an Aging America.* Washington, DC: US Government Printing Office; 1988:108–116. National Center for Health Statistics, Vital and Health Statistics Series 4, No. 25. DHHS publication no. (PHS) 89–1488, Public Health Service.
13. Chirikos TH. Aggregate economic losses from disability in the United States: a preliminary assay. *Milbank Q.* 1989;67(suppl 2, pt 1):59–91.
14. Marge M. Health promotion for persons with disabilities: moving beyond rehabilitation. *Am J Health Promotion.* 1988;2:29–35.
15. Institute of Medicine, Committee on a National Agenda for the Prevention of Disabilities; Pope AM, Tarlow AR, eds. *Disability in America: Toward a National Agenda for Prevention.* Washington, DC: National Academy Press; 1991.

Swimming-Associated Cryptosporidiosis

Frank J. Sorvillo, MPH, Kenn Fujioka, PhD, Bernard Nahlen, MD, Michael P. Tormey, MPH, Richard Kebabjian, RS, and Laurene Mascola, MD, MPH

ABSTRACT

In July and August 1988, an outbreak of gastroenteritis affected 44 of 60 (73%) persons from 5 separate swimming groups who had used the same swimming pool in Los Angeles. *Cryptosporidium* was identified in 5 of 8 (63%) stool specimens, and the clinical picture was consistent with *Cryptosporidium* infection. Resistance of *Cryptosporidium* to chlorine, an inadequately maintained pool filtration system, repeated exposure to pool water, and possible continuing pool contamination may have contributed to ongoing transmission. *Cryptosporidium* should be considered a potential etiologic agent of gastroenteritis associated with recreational water use. (*Am J Public Health.* 1992;82:742–744)

Introduction

Since 1985, *Cryptosporidium* has been increasingly recognized as an important cause of gastroenteritis.[1] Outbreaks of cryptosporidiosis have been reported among day-care center attendees,[2] among travelers,[3,4] and in communities with contaminated water supplies.[5,6] Although outbreaks of enteric infection associated with recreational water contact have been documented,[7–10] however, swimming pool–associated cryptosporidiosis has not been previously reported.

In July and August 1988, an outbreak of diarrheal illness affected members of five separate groups of swimmers using the same private pool (Pool A) in Los Angeles County. The outbreak began within 1 week after a fecal accident in the pool in early July. The five groups, who had repeated pool contact during July, were a high school water polo team, a beginning scuba class, a masters swim group, elementary school students, and the lifeguard staff. The outbreak was not reported to health authorities until early August, when an investigation was initiated.

Methods

A standardized questionnaire was used to obtain information on the basic demographics of the persons involved in the outbreak and on clinical manifestations of and exposure to potential risk factors, including contact with day-care centers, travel outside the United States, consumption of untreated surface water, animal contact, common sources of food or water, and length and type of pool contact. To assess potential risk factors, the questionnaire was administered to a cohort of 32 elementary school students who had used two local pools; 16 who had used Pool A and 16 who had used another pool, Pool B. (Three students with exposure to both pools were not included in the analysis.) In addition, a case-control study was conducted among the remaining four groups. Thirty-six ill swimmers for whom family controls were available were compared with 72 unmatched, well family members. A suspect case was defined as any swimming group member, household contact of a group member, or school classmate with watery diarrhea or diarrhea plus cramping and/or fever from July 1 through August 31. Patients whose stool tested positive for *Cryptosporidium* were considered confirmed cases.

Data from an active surveillance system were reviewed to determine whether there had been a communitywide increase of gastrointestinal illness during July and August.

Laboratory results were obtained from physicians and/or laboratories where stool specimens were examined. Additional specimens were obtained from persons who continued to be ill. Samples were assessed for bacterial (*Salmonella, Shigella, Campylobacter*) and protozoan pathogens by standard methods. Stool specimens, which had been preserved, were obtained from private laboratories and evaluated for *Cryptosporidium* by the Los Angeles County Public Health Laboratories. Auramine O stain was used as an initial screen for *Cryptosporidium*, and a modified acid-fast stain was used as a confirmatory test. Viral studies were not performed.

An assessment of pool maintenance was made; however, attempts to recover *Cryptosporidium* from water were not made because the delay in investigation diminished the likelihood of recovering oocysts.

Chi-square, the Fisher Exact Test, a chi-square test for trend, and the *t* test were used to evaluate association of possible risk factors with illness. Relative risk (RR), odds ratios (ORs), and 95% confidence intervals (CIs) were calculated where appropriate.

Frank J. Sorvillo, Kenn Fujioka, Michael P. Tormey, Richard Kebabjian, and Laurene Mascola are with Acute Communicable Disease Control of the Los Angeles County Department of Health Services. Bernard Nahlen is with the Field Services Division of the Centers for Disease Control in Atlanta.

Requests for reprints should be sent to Frank J. Sorvillo, MPH, 313 N Figueroa Street, Los Angeles, CA 90012.

This paper was submitted to the Journal January 29, 1991, and accepted with revisions September 16, 1991.

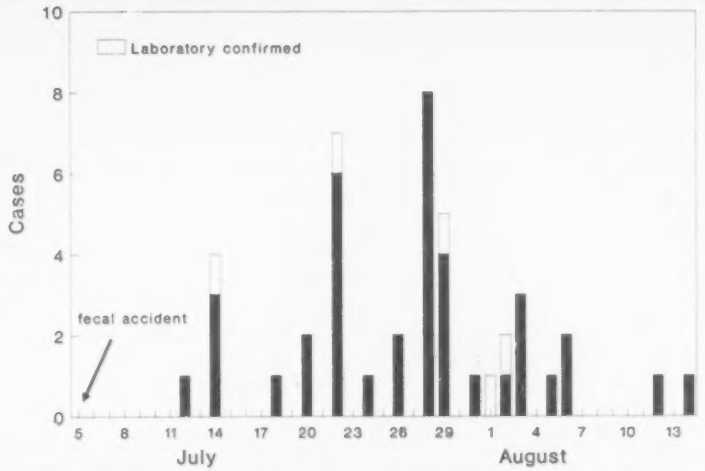


FIGURE 1—Cases of diarrheal illness among persons using Pool A, Los Angeles County, July–August, 1988.

TABLE 1—Age Range and Occurrence of Diarrhea by Group among Swimmers with Exposure to Pool A, Los Angeles County, July and August 1988

Swimming group	Median Age (y) (range)	Total Members	Interviewed (%)[a]	Ill (%)[b]
Scuba class	27 (16–34)	15	13 (87)	13 (100)
Masters group	27 (20–57)	17	13 (76)	11 (85)
Pool lifeguards	18 (14–20)	5	4 (80)	3 (75)
Water polo team	15 (13–16)	15	11 (73)	8 (73)
Elementary school group	6 (5–39)	26	19 (73)	9 (47)
Total	17 (5–57)	78	60 (77)	44 (73)

[a]Percent of total members.
[b]Percent of total interviewed.

TABLE 2—Attack Rate by Hours of Water Exposure to Pool A among Members of Five Swimming Groups, Los Angeles County, July and August 1988

Exposure in Hours[a]	Number[b]	Number Ill	%
1–3	13	5	39
4–6	8	5	63
>6	37	33	89

Note. Chi-square test for trend $P < .001$.
[a]Total hours contact during July.
[b]Two persons with undetermined length of exposure.

Results

Interviews were completed for 60 of the 78 swimming group members (77%). Between July 13 and August 14, 1988, 44 group members developed gastrointestinal illness (Figure 1). The overall attack rate was 73% (44/60) and ranged from 47% to 100% by group. The median age of patients was 17 years (range 5 to 57) (Table 1); the male:female ratio was 1.2:1. Illness was characterized by watery diarrhea (88%), abdominal cramping (86%), and fever (60%), and it was often protracted (median duration 5 days). Thirty-two percent (14/44) sought physician consultation; two patients, both from the scuba class, were hospitalized.

Among the cohort of 32 elementary school students, illness was significantly more common among the 16 using Pool A (Fisher's Exact Test, $P < .01$, RR = 8, 95% CI = 1.1, 56.8). Differences in age or gender between those with and without exposure to Pool A were not noted.

For the remaining four groups, cases (35/36) were more likely than family controls (1/72) to have a history of exposure at Pool A during the outbreak period ($P < .001$, OR = 2485, 95% CI 125,22026).

The attack rate was highest for those having heavy (more than 3 total hours) water exposure (chi-square test for trend, $P < .001$) (Table 2). Cases with fever reported twice the extent of water contact than those without fever (mean hours 23.1 and 12.2, respectively, t test $P = .09$). No other common exposures or risk factors could be identified.

Cryptosporidium was identified in stool specimens from five of eight (63%) patients tested; no other bacterial or parasitic enteric pathogens were found. Stool samples from two control patients were negative for *Cryptosporidium*.

No increase in reported diarrheal illness during July and August in the county or the community affected by the outbreak was observed. Inspection of the pool during the outbreak period revealed adequate chlorine levels (2 ppm). However, one of three diatomaceous earth filters was inoperative, and the total filtration flow rate was 30% below the 230 gallons per minute recommended.

Discussion

Several factors support the conclusion that this was a swimming pool–related outbreak of cryptosporidiosis. The epidemic curve is consistent with a continuing common source outbreak. The observed association between exposure to Pool A and illness, the increasing attack rate and severity of disease with increasing water exposure, and the temporal occurrence of the outbreak following the fecal accident implicate the pool as the most likely source of infection. Identification of *Cryptosporidium* in affected patients and the absence of other parasitic or bacterial enteric pathogens indicate that *Cryptosporidium* was the etiologic agent in this outbreak. The clinical manifestations, together with the protracted nature of the illness, are also consistent with *Cryptosporidium* infections. It is unlikely that *Cryptosporidium* was an incidental finding, given that asymptomatic infection is rare.[11] The organism is infrequently found in stool samples during nonepidemic situations.[12] *Cryptosporidium* was not a reportable disease at the time of this outbreak. Therefore, the background prevalence of *Cryptosporidium* among immunocompetent persons in Los Angeles County was unknown. In addition, the level of chlorination (2 ppm) found during inspection of the pool was sufficient to inactivate other known enteric pathogens.

Resistance of *Cryptosporidium* to chlorination,[13] an inadequately maintained filtration system, repeated and often prolonged exposure, and the apparent low infectious dose of *Cryptosporidium*[14] may have contributed to ongoing transmission in this outbreak. Continued pool use and possible ongoing contamination of the pool by infected persons, many of whom continued to swim despite their illness, also could have helped sustain transmission. The increased attack rate observed for longer water contact may be related to a greater inoculum associated with swallowed water. The scuba class, which presented a high likelihood of swallowing large quantities of water, had the highest attack rate (100%), and the only patients hospitalized were from this group. The elementary school children, who might also be expected to swallow water, had much less water exposure than the scuba group. Because attempts to recover *Cryptosporidium* from the pool water were not made and patients were not tested for viral pathogens, our findings must be interpreted with caution.

Outbreaks of *Giardia intestinalis* infection and adenovirus types 4 and 3 associated with inadequately maintained swimming pools have been reported.[8,15–17] A well-maintained, fine-grade diatomaceous earth filtration system is probably effective in removing the small *Cryptosporidium* (4–6) oocysts.[18] Whether the commonly used rapid sand filters are able to remove oocysts is unknown.

Because chlorine is ineffective against *Cryptosporidium*, closure of a pool until relatively complete filtration has occurred should be considered as an additional measure for pool disinfection following fecal contamination for pools with diatomaceous earth filtration systems. Complete draining of the pool and cleaning of filter media may be necessary for pools with rapid sand filters. Although this outbreak followed a fecal accident, reliable information on the person who defecated in the pool was unavailable, and testing was not possible. Therefore, the causal role of the fecal accident in the occurrence of the outbreak is only suggestive.

Swimming-related cryptosporidiosis may be more common than previously recognized.[19] Public health authorities should have a high index of suspicion for *Cryptosporidium* as a possible etiologic agent during outbreaks of gastroenteritis associated with recreational water contact.

Although typically a self-limited infection in the immunocompetent host,[1] *Cryptosporidium* infection, for which there is no known effective therapy, may cause an unremitting diarrhea with significant morbidity and mortality in the immunocompromised patient.[20] There are an estimated 15 000 public swimming pools and 500 000 in-ground private pools in Los Angeles County. This is also an area of relatively high AIDS incidence and HIV seroprevalence.[21] The convergence of these factors creates the potential for swimming-related exposure of *Cryptosporidium* among highly susceptible persons at high risk of potentially life-threatening illness. □

Acknowledgments

The authors wish to acknowledge the valuable technical advice of Drs. Dennis Juranek and Lawrence Ash; the laboratory expertise of Wesley Tokushige; and the field assistance of Sheila Schweid, PHN, and Ming Hilario, PHN.

References

1. Wolfson JS, Richter JM, Waldron MA, Weber DJ, McCarthy DM, Hopkins CC. Cryptosporidiosis in immunocompetent patients. *N Engl J Med.* 1985;312:1278–1282.
2. Alpert G, Bell LM, Kirkpatrick CE, et al. Outbreak of cryptosporidiosis in a day-care center. *Pediatrics.* 1986;77:152–157.
3. Jokipii L, Pohjola S, Jokipii AM. Cryptosporidiosis associated with traveling and giardiasis. *Gastroenterology.* 1985;89:838–842.
4. Waterman S, Keene W, Meade J, Sterling C, Ruiz-Matus C. Cryptosporidiosis in a group of travelers to Puerto Vallarta, Mexico, 1986. In: Steffen R, Lobel HO, Haworth J, Bradley DJ, eds. *Travel Medicine: Proceedings of the First Conference on International Travel Medicine.* Berlin, Germany: Springer-Verlag; 1989:326–327.
5. D'Antonio RG, Winn RE, Taylor JP, et al. A waterborne outbreak of cryptosporidiosis in normal hosts. *Ann Intern Med.* 1985;103:886–888.
6. Hayes EB, Matte TD, O'Brien TR, et al. Large community outbreak of cryptosporidiosis due to contamination of a filtered public water supply. *N Engl J Med.* 1989;320:1372–1376.
7. Centers for Disease Control. Human cryptosporidiosis: Alabama. *MMWR.* 1982;31:252–254.
8. Porter JD, Ragazzoni HP, Buchanon JD, Waskin HA, Juranek DD, Parkin WE. *Giardia* transmission in a swimming pool. *Am J Public Health.* 1988;78:659–662.
9. Rosenberg ML, Hazlet KK, Schaefer J, Wells JG, Pruneda RC. Shigellosis from swimming. *JAMA.* 1976;236:1849–1852.
10. Baron RC, Murphy FD, Greenberg HB, et al. Norwalk gastrointestinal illness: an outbreak associated with swimming in a recreational lake and secondary person-to-person transmission. *Am J Epidemiol.* 1982;115:163–172.
11. Casemore DP, Sands RL, Curry A. *Cryptosporidium* species a "new" human pathogen. *J Clin Pathol.* 1985;38:1321–1336.
12. Skeels MR, Sokolow R, Hubbard CV, Andrus JK, Baisch J. *Cryptosporidium* infection in Oregon public health clinic patients 1985–88: the value of statewide laboratory surveillance. *Am J Public Health.* 1990;80:305–308.
13. Campbell I, Tzipori AS, Hutchison G, Angus KW. Effect of disinfectants on survival of *Cryptosporidium* oocysts. *Vet.* 1982;111:414–415.
14. Ernest JA, Blagburn BL, Lindsay DS. Infection dynamics of *Cryptosporidium parvum* (Apicomplexa: Cryptospridiidae) in neonatal mice (*Mus muscularis*). *J Parasitol.* 1986;72:796–798.
15. Martone WJ, Hierholzer JC, Keenlyside RA, Fraser DW, D'Angelo LJ, Winkler WG. An outbreak of adenovirus type 3 disease at a private recreation center swimming pool. *Am J Epidemiol.* 1980;111:229–237.
16. D'Angelo LJ, Hierholzer JC, Keenlyside RA, Anderson LJ, Martone WJ. Pharyngoconjunctival fever caused by adenovirus type 4: report of a swimming pool-related outbreak with recovery of virus from pool water. *J Infect Dis.* 1979;140:42–47.
17. Kappus KD, Marks JS, Holman RC, et al. An outbreak of Norwalk gastroenteritis associated with swimming in a pool and secondary person-to-person transmission. *Am J Epidemiol.* 1982;116:834–839.
18. Lange KP, Bellamy WD, Hendricks DW, Logsdon GS. Diatomaceous earth filtration of *Giardia* cysts and other substances. *J Am Water Works Assoc.* 1986;78:76–84.
19. Gallaher MM, Herndon JL, Nims LJ, Sterling CR, Grabowski DJ, Hull HF. Cryptosporidiosis and surface water. *Am J Public Health.* 1989;79:39–42.
20. Soave R, Danner RL, Honig CL, et al. Cryptosporidiosis in homosexual men. *Ann Intern Med.* 1984;110:504–511.
21. Centers for Disease Control. Human immunodeficiency virus infection in the United States: a review of current knowledge. *MMWR.* 1987;36:S–6.

Newspapers: A Source for Injury Surveillance?

David Y. Rainey, MD, MPH, and Carol W. Runyan, MPH, PhD

ABSTRACT

Newspapers have not been extensively evaluated as an injury surveillance source. We compared clippings with medical examiner records for 45 residential fire deaths and 58 drownings of children to assess extent, completeness, and accuracy of newspaper coverage. Newspapers covered 96% of the fire fatalities and 78% of the drownings and contained more information than medical examiner records on several factors, including fire cause and smoke detector presence, pool fences, warning signs, and supervision of children. (*Am J Public Health.* 1992;82:745–746)

Introduction

Public health surveillance is essential for injury control. Useful data include time, place, person, injury type, causation, circumstances, medical care, and health outcomes.[1] Death certificates and medical examiner records are routinely used for mortality surveillance but may not provide sufficient information about prevention. Newspapers are an available, inexpensive source of potential information.[2,3]

This study evaluated the extent, completeness, and accuracy of coverage provided by North Carolina newspapers of childhood fatalities from fires and drownings over a 13-month period.

Methods

By state law, all fatal injuries must be investigated by a local medical examiner and the reports filed with the North Carolina Office of the Chief Medical Examiner (NCOCME). Each file contains the death certificate, a narrative report describing the event, and, sometimes, reports of autopsy, police, and toxicology investigations.

All fatalities from residential fires and drownings among persons less than 19 years of age were identified through the NCOCME for the period January 1, 1988, through January 31, 1989.

Newspaper clippings reporting injury events during the study period were obtained from a commercial clipping service collecting all 72 daily and 186 nondaily newspapers published in North Carolina. When reports were not available from the clipping service, clippings were obtained from the University of North Carolina's newspaper library, medical examiner files, and local newspapers.

All fatal fires during the 13-month period were evaluated by interviewing local fire chiefs. The variables considered are listed in Table 1.

Newspaper clippings and medical examiner reports were matched and compared for the frequency of recording of the index variables. If a variable was reported as unknown or under investigation, it was listed as "not mentioned." The unit of analysis was the decedent, even if there was more than one death per fire or drowning event. When more than one clipping was found, the combined information was used. The accuracy of newspaper reports was measured by comparing newspaper coverage and (1) the fire chief interview and the NCOCME report (for fires) and (2) the NCOCME report (for drownings).

Results

Extent of Newspaper Coverage

The NCOCME reported 45 residential fire fatalities of children under age 19 during the 13 months. Twenty-two children died in fires that claimed more than one life. The clipping service retrieved reports on 39 of the 45 children. An additional four reports were found in the university library collection, for a total of 95.6% (43/45) coverage.

The NCOCME identified 58 drownings of individuals less than 19 years of age. Seven children died in drowning incidents that claimed more than one life. The clipping service produced 42 reports of the 58 incidents. An additional 3 reports were found in the library, for a total coverage of 77.6%. No deaths were identified by the newspapers for which an NCOCME report was not available.

Completeness of Newspaper Reports

Table 1 shows that, for 43 fires and 45 drownings, some information was not well covered by either source, but the newspapers were decidedly more complete with respect to a number of the variables studied. Most notably, clippings were more likely to include information about the

This work was conducted while David Y. Rainey was at the University of North Carolina Department of Maternal and Child Health and Preventive Medicine Residency Program, and Carol W. Runyan with the Department of Health Behavior and Health Education and Injury Prevention Research Center, University of North Carolina, Chapel Hill. David Y. Rainey is now with the Department of Pediatrics, University of Colorado.

Requests for reprints should be sent to Carol W. Runyan, MPH, PhD, Director, Injury Prevention Research Center, Rosenau Hall, CB 7400, Chapel Hill, NC 27599.

This paper was submitted to the Journal March 19, 1991, and accepted with revisions July 18, 1991.

cause of the fire or presence of a smoke detector than was the NCOCME report. However, medical examiner reports were much more likely to include the cause of death, alcohol use by the decedent, and race. Race was never reported directly in the newspaper but could be inferred from photographs in 15% of cases.

Findings were similar for drownings. The NCOCME report was more complete on race of the decedent, alcohol involvement, purported swimming ability, and resuscitation attempts. However, newspapers included more information about child supervision and characteristics of the body of water (public ownership, warning signs, and fencing).

Accuracy of Newspaper Reports

Newspapers and the fire chief interview agreed completely on age, sex, and race of the decedents and type of construction materials, but agreement was only 90.6% (39/43) for cause of the fire. The four disagreements resulted from the newspaper reporting a specific cause while the fire chief stated that the cause was unknown. Newspapers and medical examiner reports agreed completely on age, sex, and race of the decedent and type of water body.

Discussion

This study suggests that newspapers may be a useful adjunct to other injury surveillance efforts. Newspapers were more complete than NCOCME reports in supplying information about the cause and circumstances of fires and presence of protective factors for both fires and drownings. On the other hand, newspapers were not good sources of information about race or the possible involvement of alcohol. Otherwise, newspapers were accurate in reporting demographic information and causes and circumstances of fires.

Two issues are of particular interest. First, our results demonstrate that for childhood fire deaths and drownings, newspapers reported 96% and 78% of the cases, respectively. However, we located these articles knowing from the NCOCME that a death had occurred. To be optimally useful for surveillance, newspapers should be able to identify cases independently. Relying only on the clipping service would have resulted in reports of 91% of the fire deaths and 72% of the drownings.

Second, this analysis revealed that there is information about protective factors that is more likely to appear in the newspaper than in official records. This finding supports the value of newspapers and signals an area in which the medical examiner system might improve documentation. In fact, items to indicate the use of life preservers, swimming ability, and the presence of smoke detectors recently have been added to the NCOCME reporting form.

Besides surveillance, the reporting of injury events has potential value in raising public awareness about injury risks and protective measures. Newspapers should be encouraged to cover injury events fully and provide detailed information about protective measures. Injury surveillance efforts should consider newspapers as a potential source of valuable supplementary information, while still focusing on improving other systems of record keeping.

This study did not attempt to assess the specific methods or costs of the clipping service per se, factors that may vary by setting. Further research could document whether these findings generalize to other less dramatic or newsworthy injury problems and whether there are systematic differences in reporting policies or practices by publisher or locale. □

TABLE 1—Completeness of Coverage in Medical Examiner Reports and Newspapers: Child Fatalities in Residential Fires (n = 43) and Drownings (n = 45), North Carolina

	Medical Examiner (times reported, %)	Newspaper (times recorded, %)
Fires		
Age	100.0	100.0
Sex	100.0	100.0
Race	100.0	16.3
Time of drowning	100.0	100.0
Cause of death	100.0	53.5
Supervision of children	69.8	88.4
Home construction material	44.2	58.1
How the fire was discovered	37.8	67.4
Cause of fire	30.2	62.6
Alcohol involvement	23.2	0
Time to intervention	14.0	27.9
How fire department was contacted	11.6	44.2
Presence of smoke alarm	4.6	18.6
Presence of fire extinguisher	0	2.3
Drownings		
Age	100.0	100.0
Sex	100.0	100.0
Race	100.0	8.9
Location of drowning	100.0	100.0
Time of drowning	100.0	100.0
Type of water body	100.0	100.0
Alcohol presence	84.4	0
Time submerged	71.1	71.1
CPR attempt	66.6	33.3
Supervision of children (<age 16, n = 29)	65.5	79.3
Length of rescue attempt	33.1	27.7
Ability to swim (age > 3, n = 34)	25.0	13.9
Private/public property	4.4	24.4
Warning sign presence	2.2	22.2
Water area fenced in	0	18.2

Note. Percentages are based only on the cases for which both a newspaper clipping and the medical examiner record were available. CPR = cardiopulmonary resuscitation.

Acknowledgments

This work was supported, in part, by the University of North Carolina Injury Prevention Research Center, with funds from the Centers for Disease Control (CCR40222-04). Dr. Rainey's time was supported by the University of North Carolina Preventive Medicine Residency Program.

The authors wish to acknowledge the conceptual, editorial, and technical assistance of Dr. Shrikant Bangdiwala, Dr. Jane Brown, Dr. John Butts, Dr. Thomas Cole, Ms. Janet Heath, Ms. Mary Linzer, Dr. Lewis Margolis, Ms. Nancy Moore, Mr. Michael Patetta, Dr. Carol Reuss, Dr. Desmond Runyan, Dr. Jeffrey Sacks, and Dr. Earl Siegel.

References

1. Committee on Trauma Research. *Injury in America.* Washington, DC: National Academy Press; 1985.
2. Murphy DJ, Seltzer BL, Yesalis CE. Comparison of two methodologies to measure agriculture occupational fatalities. *Am J Public Health.* 1990;80:198–200.
3. Sacks JJ, Sattin RW, Bonzo SE. Dog bite related fatalities from 1979–1988. *JAMA.* 1989;262:1489–1492.

Response Comparability of Family and Staff Proxies for Nursing Home Residents

Sharon L. Tennstedt, PhD, Katherine M. Skinner, PhD, Lisa M. Sullivan, PhD, and John B. McKinlay, PhD

ABSTRACT

Response comparability on a wide variety of factual and judgmental information regarding nursing home residents was assessed for two proxy groups: family and staff. Family proxies had the most information. Family and staff proxies are not interchangeable. A general lack of response comparability suggests that proxies for nursing home residents should be selected in relation to the type of data needed and the time period of interest. (*Am J Public Health.* 1992;82:747–749)

Introduction

Because of the high prevalence of impairments among nursing home residents, proxy respondents are frequently required. Most studies of proxy response comparability have examined agreement between subject and proxy on factual information such as occupation, diet, smoking, and other disease risk factors,[1-3] although some have also examined agreement on items requiring judgment of health and functional status.[4,5] Much of this work has not been done on older populations, however; nor has it included nursing home residents. Further, most studies have not investigated the response comparability of two proxies of differing relationships—e.g., staff or family—to the elder subject. To examine the question of *who* is the best proxy respondent in studies of nursing home residents, this study (1) investigated proxy-proxy agreement on items involving both fact and judgment, and (2) examined the extent to which the subject-proxy relationship affected the extent and type of knowledge about the subject.

Methods

This study was part of a longitudinal study of the functional status and receipt of care of a geographically stratified sample of 4774 adults aged 70 or older, randomly selected from local census lists in Massachusetts. At first follow-up in 1989, the first 47 respondents who were determined to be alive and were admitted to a nursing home since the baseline survey 4 years earlier were identified for this study. Using the same instrument, interviews were conducted with these nursing home residents and two different proxies—i.e., a staff member familiar with the patient, as selected by the facility, and a family member, usually the preadmission caregiver. Order of interviews was assigned randomly. Resident and staff interviews were conducted in person at the same visit. Family interviews were conducted by telephone to contain costs.

Data regarding the resident included sociodemographics, date and reason for admission, preadmission use of formal services, preadmission and current receipt of informal care, and functional limitations. The Short Portable Mental Status Questionnaire (SPMSQ)[6] was administered to the resident to assess cognitive status. Additional data collected from the two proxy groups included frequency of current contact with the resident and current provision of care; knowledge of the resident's functional abilities; and, for family proxies, preadmission residential and caregiving status in relation to the resident. Judgment about the resident's functional limitations was reached through five questions designed to classify nursing home patients functionally.[7] These questions, their coding, and the summary statistics are provided in Table 1.

The percent agreement and Cohen's kappa (κ) statistic were calculated for categorical information provided by both proxies. Ranging from less than 0 to 1, κ values greater than 0.75 indicate excellent agreement and κ values less than 0.40 indicate poor agreement.[8] To test the significance of the κ statistic (i.e., the level of agreement that is more than expected by chance), z scores were computed from the observed κ statistics and their standard errors. One-tailed significance levels to detect more, rather than less, agreement by chance were also computed.

Results

Sample Description

The nursing home residents were predominantly female (70.2%), not married (91.5%), and very old (68.1% older

The authors are with the New England Research Institute in Watertown, MA.

Requests for reprints should be sent to Sharon Tennstedt, PhD, New England Research Institute, 9 Galan Street, Watertown, MA 02172.

This paper was submitted to the Journal March 6, 1991, and accepted with revisions September 20, 1991.

TABLE 1—Judgment of Functional Limitations

Question	Coding	Family Proxies		Staff Proxies	
		Mean	SD	Mean	SD
Does (resident) have any problem walking?	1. is independent 2. uses equipment and/or device 3. needs assistance 4. needs assistance *and* equipment/device 5. does not ambulate	2.87	1.71	2.85	1.56
Does (resident) have any problem dressing?	1. is independent 2. uses equipment and/or device 3. needs supervision 4. needs assistance 5. is totally dependent	3.72	2.01	3.70	1.38
Does (resident) need help with eating?	1. is independent 2. uses adaptive utensils 3. needs supervision 4. needs assistance 5. must be fed; needs tube/ parenteral feedings	2.51	2.04	2.23	1.40
Does (resident) have any bladder accidents?	1. is continent 2. uses catheter or ileal loop—self care 3. is incontinent <3x/wk 4. is incontinent ≥3x/wk 5. needs catheter or ileal loop—staff care	3.36	2.38	2.66	1.70
Does (resident) have any problem getting to or using the toilet?	1. is independent 2. needs supervision 3. needs assistance getting to/from toilet 4. needs assistance in use of bedpan/toilet 5. is totally dependent	3.04	2.47	2.70	1.90

than 85 years). The vast majority (82.5%) were scored as having moderate cognitive impairment on the SPMSQ (i.e., a score of 5 or more, adjusted for education) and were unable to complete the interview. Most (74.5%) had been admitted in the previous 2 years.

Most of the family proxies were offspring (63.8%), usually daughters, followed by other relatives (31.9%) and nonkin guardians (4.3%). Although only 31.9% had lived with the resident prior to admission, 78.7% had provided informal care. Staff proxies were social workers (72.3%) or nurses (27.7%). Frequency of contact was greater for staff, but 57.4% of family proxies saw the resident at least several times a week. Self-rated knowledge of the resident was equally good for both proxy groups.

Response Comparability

There was good to excellent agreement between the two proxies on the resident's sociodemographic characteristics (see Table 2). Agreement declined for questions related to the preadmission period. Although there was good agreement on who provided informal care, agreement was not good on types of informal care and formal services used. Staff proxies responded "don't know" more frequently than family proxies. Similarly, there was good agreement on who was currently involved with the resident but not on what the proxy did, except for providing transportation. There was generally poor agreement on most questions relating to major health problems except for diabetes, cancer, and heart conditions. Finally, there was slightly better agreement on items regarding functional limitations. Surprisingly, the area of least agreement was the ability of residents to feed themselves.

Discussion

The high prevalence of cognitive impairment in this sample of nursing home residents clearly suggests a need for proxy respondents in surveys of such populations. However, because most of the sample had moderate to severe cognitive impairment, it was not possible to ascertain response comparability between each of the proxies and the resident subjects themselves. Yet this study advances our knowledge by examining response comparability on general survey items involving both facts and judgment as well as between two different sources of proxy information.

It appears that family members have sufficient knowledge about the resident's current status and know more about the resident's preadmission data to be informed proxies. However, while item nonresponse for family member proxies is lower than it is for staff proxies, it is not possible to validate family proxy responses without self-report or medical record data. Such responses provide more information, but their accuracy requires further investigation. Despite this limitation, these findings have implications for study methodology.

The data suggest that proxies should be selected in relation to both the type of data needed and the time period of interest. Further, proxy selection may have to be item specific, and more than one proxy may be required to gather complete data.

An area requiring further study involves judgment of functional status. Other studies[4,5] have reported that proxies underestimate functional abilities when compared with elders' self-reports. This study showed lack of agreement between the two proxies. Perhaps there is no substitute for clinical observation of functional behaviors when collecting this type of data.

Finally, there is the possibility that the differing mode of administration (tele-

TABLE 2—Response Comparability: Family vs Staff Proxies

	Kappa	% Agree	% Don't Know: Family	% Don't Know: Staff
Resident's sociodemographic characteristics				
Marital status	.849*	93.6	—	—
Education	.456*	70.2	10.6	23.4
Religion	.915*	95.7	—	2.1
Preadmission status				
Prior household composition	.393*	56.5	2.1	21.3
Reason for admission	.382*	57.4	—	—
Overnight hospitalization	.283	65.9	2.1	21.3
Relationship of primary caregiver	.616*	72.7	—	9.1
Areas of informal care[a]				
Housekeeping	.190	39.3	—	35.7
Meals	.196	39.3	—	39.3
Personal care	.273*	42.9	—	42.9
Shopping	.194*	46.4	—	39.3
Medications	.077	28.6	3.6	42.9
Transportation	.119	35.7	—	39.3
Areas of formal service[a]				
Homemaker	.097	35.7	—	35.7
Meals	.029	32.1	—	39.3
Personal care	−.007	25.0	—	42.9
Shopping	.024	42.9	—	39.3
Transportation	−.008	39.3	—	39.3
Medication	−.018	35.7	3.6	42.9
Current informal care				
Relationship of primary caregiver	.711*	81.1	—	19.1
Areas of Care				
Personal care	.000	87.2	2.1	—
Meals	.305	63.8	—	2.1
Shopping	.183	68.1	—	—
Transportation	.405	70.2	—	2.1
Finances	.251	70.2	—	12.8
Laundry	.330*	70.2	—	12.8
Health conditions				
Diabetes	.798*	95.7	—	—
Cancer	.000	95.7	—	—
Arthritis	.141	44.7	8.5	2.1
Heart condition	.417*	63.8	—	2.1
Hypertension	.297	74.5	4.3	—
Cardiovascular accident	.245	61.7	2.1	4.3
Mental disorder	.280	53.2	—	—
Functional limitations				
Walking	.583*	68.1	—	—
Dressing	.447*	59.6	4.3	—
Eating	.148	44.7	4.3	—
Bladder accidents	.339*	53.2	10.6	2.1
Toileting	.394*	57.5	8.5	2.1
Wears glasses	.363	68.0	—	6.4
Uses hearing aid	.545*	85.1	2.1	—
Moves wheelchair—length of room[b]	.778*	87.5	12.5	—

*$P \leq .001$; selected as criterion of significance to ensure 96% likelihood of no type I error.
[a]n = 28; questions not asked for those residents hospitalized for month prior to admission.
[b]n = 16.

phone vs in-person interview) influenced the survey results. However, given the results of work by others[9–12] and the nonsensitive nature of the data collected, it is unlikely that response differences resulted from mode effects. □

Acknowledgments

This research was supported by National Institute on Aging Grant AG07182.

A version of this paper was presented at the 117th Annual Meeting of the American Public Health Association in October 1989.

The authors acknowledge the helpful comments and input of Sybil Crawford, Henry Feldman, and Alan Jette.

References

1. Humble CG, Samet JM, Skipper BE. Comparison of self- and surrogate-reported dietary information. *Am J Epidemiol.* 1984;119:86–98.
2. Herrmann N. Retrospective information from questionnaires: I. comparability of primary respondents and their next-of-kin. *Am J Epidemiol.* 1985;121:937–947.
3. Lerchen ML, Samet JM. An assessment of the validity of questionnaire responses provided by a surviving spouse. *Am J Epidemiol.* 1986;123:481–489.
4. Magaziner J, Simonsick E, Kashner TM, Hebel JR. Patient-proxy response comparability on measures of patient health and functional status. *J Clin Epidemiol.* 1988;41:1065–1074.
5. Rubenstein LZ, Schairer C, Weiland GD, et al. Systematic biases in functional status assessment of elderly adults: effects of different data sources. *J Gerontol.* 1984;39:686–691.
6. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *J Am Geriatr Soc.* 1975;23:433–441.
7. Morris JN, Sherwood S, May MI, Bernstein E. FRED: an innovative approach to nursing home level-of-care assignments. *Health Serv Res.* 1987;22:117–138.
8. Fleiss JL. *Statistical Methods for Rates and Proportions.* New York, NY: John Wiley & Sons; 1973.
9. Kulka RA, Weeks RA, Lessler JT, Whitmore RW. A comparison of the telephone and personal interview modes for conducting local household health surveys. In: National Center for Health Services Research, Health Survey Research Methods. *Proceedings of the Fourth Conference on Health Survey Research Methods.* NCHSR Proceedings Series. Washington, DC: US Government Printing Office; 1984. US Dept of Health and Human Services publication PHS 84–33466.
10. Miller PV. A comparison of the telephone and personal interviews in the Health Interview Survey. In: National Center for Health Services Research, Health Survey Research Methods. *Proceedings of the Fourth Conference on Health Survey Research Methods.* NCHSR Proceedings Series. Washington, DC: US Govt Printing Office; 1984. US Dept of Health and Human Services publication PHS 84–33466.
11. Siemiatycki J. A comparison of mail, telephone, and home interview strategies for household health surveys. *Am J Public Health.* 1979;69:238–245.
12. Weeks MF, Kulka RA, Lessler JT, Whitmore RW. Personal versus telephone surveys for collecting household health data at the local level. *Am J Public Health.* 1983;73:1389–1394.

A26234

American Journal of Public Health

June 1992, Vol. 82, No. 6 Established 1911

Editorial

785 Community Health, Community Risks, Community Action *S. Shea*

Public Health Policy Forum

788 Editorial: Managed Care for the Seriously Mentally Ill *D. Mechanic*

790 Use of Community-Based Mental Health Programs by HMOs: Evidence from a Medicaid Demonstration *J. B. Christianson, N. Lurie, M. Finch, I. S. Moscovice, and D. Hartley*

796 Commentary: Caring for the Indigent Mentally Ill—New Strategies and Old Problems *E. Ferran, Jr*

Featuring Health in the Community

799 Weight Gain Prevention and Smoking Cessation: Cautionary Findings *S. M. Hall, C. D. Tunstall, K. L. Vila, and J. Duffy*

804 Improving Referral Compliance after Public Cholesterol Screening *L. A. Maiman, N. G. Hildreth, C. Cox, and P. Greenland*

810 Food Sources, Dietary Behavior, and the Saturated Fat Intake of Latino Children *C. E. Basch, S. Shea, and P. Zybert*

816 Socioeconomic Status and Health: How Education, Income, and Occupation Contribute to Risk Factors for Cardiovascular Disease *M. A. Winkleby, D. E. Jatulis, E. Frank, and S. P. Fortmann*

821 Socioeconomic and Behavioral Correlates of Body Mass Index in Black Adults: The Pitt County Study *J. B. Croft, D. S. Strogatz, S. A. James, N. L. Keenan, A. S. Ammerman, A. M. Malarcher, and P. S. Haines*

827 Prevention of Cigarette Smoking through Mass Media Intervention and School Programs *B. S. Flynn, J. K. Worden, R. H. Secker-Walker, G. J. Badger, B. M. Geller, and M. C. Costanza*

835 The Second Chicago Televised Smoking Cessation Program: A 24-Month Follow-up *R. B. Warnecke, P. Langenberg, S. C. Wong, B. R. Flay, and T. D. Cook*

841 Functional Limitations and Disability among Elders in the Framingham Study *M. Kelly-Hayes, A. M. Jette, P. A. Wolf, R. B. D'Agostino, and P. M. Odell*

846 Neuroleptic Medication and Prescription Practices with Sheltered-Care Residents: A 12-Year Perspective *S. P. Segal, D. Cohen, and S. R. Marder*

853 Eligibility for Publicly Financed Home Care *M. E. Jackson, B. Burwell, R. F. Clark, and M. Harahan*

867 First-Year Impact of the 1989 California Cigarette Tax Increase on Cigarette Consumption *R. L. Flewelling, E. Kenney, J.P. Elder, J. Pierce, M. Johnson, and D. G. Bal*

870 Drug Use and Suicide Ideation and Behavior among North Carolina Public School Students *W. M. Felts, T. Chenier, and R. Barnes*

873 Delay in Ambulance Dispatch to Road Accidents *H. Brodsky*

875 Race, Belief in Destiny, and Seat Belt Usage: A Pilot Study *I. Colón*

(continued page 777)

ED KIANKA
CONGRESSIONAL INFORMATION SVC
4520 EAST-WEST HIGHWAY
BETHESDA MD 20814
86389
028

APHA
American Public Health Association

Featuring Health in the Community

Contents

American Journal of Public Health

The Journal's one-page guidelines for authors, "What *AJPH* Authors Should Know," are printed in each issue. They can also be obtained from the Editorial Office.

June 1992, Vol. 82, No. 6 ISSN: 0090-0036 Established 1911

877 Work-Site Nutrition Intervention and Employees' Dietary Habits: The Treatwell Program *G. Sorensen, D. M. Morris, M. K. Hunt, J. R. Hebert, D. R. Harris, A. Stoddard, and J. K. Ockene*

881 Health Status of Older Male Prisoners: A Comprehensive Survey *P. L. Colsher, R. B. Wallace, P. L. Loeffelholz, and M. Sales*

885 A Neonatal Hepatitis B Surveillance and Vaccination Program: New York City, 1987 to 1988 *K. J. Henning, D. M. Pollack, and S. M. Friedman*

Articles

857 Postmarketing Surveillance Study of Permethrin Creme Rinse *E. B. Andrews, M. C. Joseph, M. J. Magenheim, H. H. Tilson, P. A. Doi, and M. W. Schultz*

862 Do Time Trends in Food Supply Levels of Macronutrients Reflect Survey Estimates of Macronutrient Intake? *N. T. Crane, C. J. Lewis, and E. A. Yetley*

Public Health Briefs

889 Physician and Infection Control Practitioner HIV/AIDS Reporting Characteristics *J. L. Jones, P. Meyer, C. Garrison, L. Kettinger, and P. Hermann*

Notes from the Field

892 Falls among the Elderly: A Community Prevention Program *R. Tideiksaar*

893 Community-Based Case Management of HIV Disease *F. J. Payne, C. S. Sharrett, D. N. Poretz, L. J. Eron, T. Stage, R. Foroobar, C. Bowman, and R. K. Miller*

Other Journal Departments

778 Association Masthead

782 Advertisers' Index

784 What *AJPH* Authors Should Know

785 Journal Masthead

798 Erratum

895 Letters to the Editor

902 Book Corner

911 Job Opportunities

American Public Health Association

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.

APHA OFFICERS
Joyce C. Lashof, MD, *President*
C. William Keck, MD, MPH, *Immediate Past President*
Helen Rodriguez-Trias, MD, *President-Elect*
Arnold E. Greenberg, SM, *Vice-President (USA)*
Trevor Hancock, MB, MHSc, *Vice-President (Canada)*
Julio Teja Perez, MD, *Vice-President (Latin America)*
Caswell A. Evans, Jr, DDS, MPH, *Treasurer*
Frances M. Veverka, RS, MPH, *Chair, Executive Board*
George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*
William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD
Frances M. Veverka, RS, MPH (1992) *Chair*
Marilyn Aguirre-Molina, EdD, MS (1993)
Brian A. Dobrow, MPH (1992)
Caswell A. Evans, Jr, DDS, MPH, *Treasurer*
Kristine M. Gebbie, BSN, MN (1994) *Vice-Chair*
Alan R. Hinman, MD, MPH (1995)
Janet Heinrich, DrPH, MPH (1993)
Margaret H. Jordan, RN, MPH (1992)
C. William Keck, MD, MPH, *Immediate Past President*
Margo D. Kerrigan, MPH (1995)
Joyce C. Lashof, MD, *President*
Barry S. Levy, MD, MPH (1993)
Helen Rodriguez-Trias, MD, *President-Elect*
Allan G. Rosenfield, MD (1994)
Reed V. Tuckson, MD (1994)
Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO
William H. McBeath, MD, MPH, *Executive Director*
Joanne E. Lukomnik, MD, *Chair, Action Board*
George Pickett, MD, MPH, *Chair, Program Development Board*
George E. Hardy, MD, MPH. *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc., 1015 Fifteenth St., NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50,000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health,* 1015 Fifteenth St., NW, Washington, DC 20005.

Information on Manuscript Preparation is published in this Journal in January. The Journal publishes reports of original research, demonstrations, evaluations, and other articles covering current aspects of public health. Manuscripts are accepted for consideration with the understanding that they, or their essential substance, have not been published previously or accepted for publication and are not under consideration elsewhere.

Subscription price, individual, is $80/year US, $120/year international; **institutional,** $160/year US, $240/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to American Journal of Public Health, 1015 Fifteenth St., NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact: Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact: Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St., NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols. 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS Bibliographic Retrieval Service Inc, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc., Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Use of Community-Based Mental Health Programs by HMOs: Evidence from a Medicaid Demonstration

Jon B. Christianson, PhD, Nicole Lurie, MD, MSPH, Michael Finch, PhD, Ira S. Moscovice, PhD, and David Hartley, MHA

ABSTRACT

Background. Proposals to enroll Medicaid beneficiaries in health maintenance organizations (HMOs) have raised concerns that community-based mental health treatment programs would be adversely affected.

Methods. In Hennepin County (Minnesota) 35% of Medicaid beneficiaries were randomly assigned to prepaid plans. Random samples of individuals with severe mental illness were selected from the prepaid enrollees and from beneficiaries remaining with traditional Medicaid. The two groups were compared with respect to their use of community treatment programs and the write-off (the proportion of patient charges for which payment was not received) experienced by those programs for members of the study sample.

Results. There was no strong evidence that Medicaid beneficiaries with severe mental illness who were randomly assigned to prepaid plans used community-based mental health treatment programs differently than did other Medicaid beneficiaries. However, write-offs were consistently higher for enrollees in prepaid plans.

Conclusions. In the short run, the use of community-based mental health treatment programs need not be affected by enrollment of Medicaid beneficiaries in prepaid plans, providing that Medicaid program administrators take steps to minimize the disruption of ongoing treatment, offer beneficiaries a choice among prepaid plans, and encourage community treatment programs to contract with plans to serve beneficiaries. (*Am J Public Health.* 1992;82:790–796)

Introduction

Proposals to enroll people with severe mental illness, particularly Medicaid beneficiaries, in health maintenance organizations (HMOs) have generated concerns on the part of mental health policy analysts and providers.[1–4] Among these concerns is that beneficiaries enrolled in HMOs will use the mental health treatment programs available in their communities less frequently than do beneficiaries under traditional Medicaid. It is feared that this reduced use, particularly if accompanied by reduced payments to treatment programs by HMOs, will have an adverse effect on the financial viability of these programs. In this article we analyze use and reimbursement of community-based mental health treatment programs on the part of Medicaid beneficiaries with severe mental illness. The experience of beneficiaries randomly assigned to prepaid plans is contrasted with that of beneficiaries in traditional Medicaid. Throughout the article, we use the term *community-based treatment program* to include community mental health centers, crisis centers, day treatment programs, and drop-in centers, all of which are typically major providers of mental health care to indigent populations.

Concerns about the Interface between HMOs and Community-Based Treatment Programs

The manner in which HMOs relate, or might relate, to community-based mental health treatment programs has been a matter of debate.[5,6] A recent survey of the benefits provided by HMOs to their private sector enrollees reported that 53% exclude treatment of chronic mental illness from their standard benefit plan and 64% exclude long-term psychotherapy.[7] These figures suggest that many HMOs may have limited experience in delivering services to severely mentally ill persons and have lead to concerns that HMOs may not take full advantage of community-based programs in treating their enrollees. However, HMO enrollees with severe mental illness may seek out publicly subsidized community-based programs to avoid HMO gatekeepers or to circumvent other HMO policies that restrict access. Such a strategy seems particularly likely for Medicaid beneficiaries, who may be familiar with these programs through past use. In some cases, HMOs may encourage the use of community-based treatment programs through contractual and other, less formal, methods, to reduce their own costs of treating enrollees with severe mental illness.[8,9] An HMO would realize cost reductions if the charges of the community programs were less than the cost of delivering similar services through HMO providers or if the HMO did not fully reimburse the charges billed by these treatment programs.

If the overall use of community treatment programs drops as a result of enrolling Medicaid beneficiaries in HMOs, program supporters fear that public sector subsidies and Medicaid reimbursements will be reduced. Such reductions could force these programs to scale down their activities or, in the extreme, to discontinue their operations. However, program supporters also foresee possible negative consequences, from a public policy standpoint, of *greater* use of community-based programs by HMO enrollees. For example, the public sector might, in effect, pay twice for the mental health care of HMO enrollees: once through general public subsidies provided to the programs and once through the capitation payment made by Medicaid to the HMO, which is intended to

The authors are with the Division of Health Services Research and Policy, School of Public Health, University of Minnesota, Minneapolis. Nicole Lurie is also with the Department of Medicine, Hennepin County Medical Center, Minneapolis, Minn.

Requests for reprints should be sent to Jon B. Christianson, PhD, Division of Health Research Services and Policy, School of Public Health, University of Minnesota, 420 Delaware Street, SE, Box 729, Minneapolis, MN 55455.

This paper was submitted to the Journal September 30, 1991, and accepted with revisions March 12, 1992.

cover all of the costs of providing care to HMO enrollees. If the programs were reimbursed appropriately by HMOs for services they provided, then, at least in theory, public subsidies to these programs could be reduced so that double paying did not occur and the programs would not be adversely affected. On the other hand, if budget subsidies were reduced by amounts greater than the reimbursements received by community-based programs for treating HMO enrollees, these programs would be faced with the need to provide more services but with fewer resources.

We tested two general hypotheses related to these issues. The first hypothesis was that the use of community-based treatment programs by severely mentally ill Medicaid beneficiaries would differ from the use observed under traditional Medicaid. Given the conflicting arguments that can be made, we did not assign a directional expectation to this hypothesis. Also, we did not address whether reductions in the use of community-based treatment programs would imply poorer quality care or result in worse patient outcomes, even though these are obviously important research issues. The second hypothesis was that the proportion of patient charges for which payment is not received by community-based programs (which we define as the "write-off") would be higher for Medicaid beneficiaries in HMOs than for beneficiaries in traditional Medicaid.

The Hennepin County Demonstration

In 1982, the Health Care Financing Administration authorized six states to enroll Medicaid beneficiaries in prepaid health plans on a demonstration basis. As part of this demonstration, in Hennepin County (Minneapolis), Minnesota, 35% of all Medicaid beneficiaries were randomly assigned to receive services from prepaid plans. They were then provided with assistance in selecting a plan; those who did not make a choice ultimately were assigned to a plan. Enrollment of individuals who were disabled owing to mental illness occurred between November 1986 and April 1987; service delivery for early enrollees began on January 1, 1987.[10]

The four prepaid plans that enrolled beneficiaries in the disabled category were all individual practice association plans. Their sponsoring organizations were Blue Cross/Blue Shield, Hennepin County, the University of Minnesota, and an independent organization affiliated with a local hospital corporation. The Hennepin County plan subcontracted with the county hospital (a major provider of mental health services) and county-subsidized community-based treatment programs for the delivery of mental health care. Blue Cross/Blue Shield also offered access to a broad array of mental health providers, including many community-based programs.

The prepaid plans were reimbursed at 95% of estimated Medicaid costs (projected on the basis of historical trends), with different rates established on the basis of age, sex, Medicare participation, institutional vs noninstitutional residence, and eligibility status. The plans were required to deliver all services covered under traditional Medicaid, including all necessary mental health care; were screened to ensure that they had the ability to provide the required range of mental health services; and were subsequently monitored by the state to ensure that they delivered these services effectively.[11] As part of their participation in the demonstration, they were expected to refer enrollees, as needed, to community-based treatment programs and to reimburse these programs for their services.

In implementing the demonstration, several steps were taken to address the concerns expressed by mental health advocates and providers that the prepaid plans would not be able to develop effective linkages with community-based treatment programs and therefore that the continuity of treatment for patients would be interrupted.[12] For instance, the demonstration administrator in Hennepin County had experience in mental health care, as did staff members in the state demonstration office. Also, the availability of a county-sponsored plan gave beneficiaries an option that provided access to most community-based treatment programs. In addition, enrollees were permitted to switch plans within 60 days of their initial choice, giving them a second chance to select plans that allowed them to maintain their relationships with mental health treatment providers. To minimize transition-related problems, an informal grievance procedure was developed that, in most cases, resulted in the prepaid plan's agreeing to pay for completion of treatment by the original provider or completion of the recommended course of treatment by plan providers. A demonstration advisory group was formed to develop standards for monitoring quality of care, for conducting medical records review, and for managing the special needs of mentally ill beneficiaries. At the end of the first demonstration year, plan contracts were amended to enable enrollees to self-refer to mental health providers without authorization from a family practice physician or a psychiatrist.[13]

At the end of the second program year, which was the first year in which disabled beneficiaries were enrolled in prepaid plans, the Blue Cross/Blue Shield plan withdrew from the demonstration, citing unexpectedly high use of services by plan enrollees. Because of Medicaid administrators' concern about the capacity and willingness of the remaining three plans to assimilate all of Blue Cross' disabled enrollees, these individuals were transferred back to traditional Medicaid, although the overall demonstration continued.

Data and Methods

Study Sample

First, Medicaid recipients aged 18 to 65 years who were classified as disabled were identified from Medicaid records. Then an algorithm based on ICD-9 diagnosis codes and the number and frequency of claims was developed to identify beneficiaries in this group with severe mental illness.[10] Applying this algorithm to Medicaid claims tapes for the 2 years before November 1986 yielded 500 individuals who had been randomly assigned to the prepaid group. However, 104 were excluded from the study for a variety of reasons (e.g., spoke a language other than English, were deceased, had moved out of area). Completed in-person baseline interviews were obtained for 93% of the potential interviewees (369 individuals) in the prepaid group. In a similar fashion, 510 individuals remaining in traditional Medicaid were identified as severely mentally ill, 90 were excluded from the interview sample, and 93% (370) of the remainder were interviewed at baseline.

Follow-up interviews were planned for 1 year after the baseline interview, but when the state discontinued participation on the part of the disabled group effective January 1988, modification of the interview schedule was necessary. For individuals enrolled in prepaid plans, follow-up interviews were administered during the period between notification of the state's intent to transfer them back to traditional Medicaid and 2 weeks following the transfer. Because intake dates for enrollment, and therefore entry into the study, varied, the period between baseline and follow-up interview ranged from 7 to

TABLE 1—Characteristics of Sample at Baseline

	Prepaid	Traditional Medicaid	*P* Value
Demographics			
Age, y	41.5	41.6	.91
Single, %	95.5	94.9	.12
Education, y	12.1	11.7	.10
Income, $/month	384	404	.31
Female, %	52.8	58.9	.12
Residence in Henepin County, y	21.9	22.9	.40
White, %	84.3	83.9	.97
Health			
Self-rated general health status index (range, 4–16)	10.2	10.1	.53
Physical functioning index (range, 0–9)	4.2	4.6	.11
Global assessment scale	55.1	55.0	.92
Social and community function			
Social contact index (range, 4–24)	13.0	12.7	.39
Assaulted, robbed, or raped in previous year, %	13.8	17.5	.21
Suicide attempt in previous year, %	6.2	9.6	.13
Living in sheltered setting, %	29.3	29.4	.99
Utilization of inpatient care			
Physical health admission in past year, %	20.4	25.4	.13
Mental health admission in past year, %	23.6	22.9	.90
Chemical dependency admission in past year, %	4.5	6.8	.26
Utilization of outpatient care			
Physical health visit in past 3 months, %	66.6	71.4	.19
Mental Health visit in past 3 months, %	77.4	73.1	.21

12 months, with an average length of approximately 11 months. Follow-up data for the comparison group were collected according to the same schedule, with individuals randomly assigned interview dates from 7 to 12 months after their baseline interviews, matching the prepaid group on a one-to-one basis. Follow-up interviews were completed with 354 individuals in the prepaid group and 366 in the traditional Medicaid group, resulting in complete baseline and follow-up data for about 96% of the individuals who completed the baseline survey. As the comparisons in Table 1 indicate, there were no significant differences in demographic characteristics or general use of health services between the two groups.

A second data source was used to supplement the self-report data collected in the in-person interviews. At the follow-up interview, respondents were asked to identify community-based treatment programs they had used since the date of the baseline interview. Record abstractors then visited these programs, beginning approximately 6 months after the date of the follow-up interview to ensure completeness of billing records, and recorded all charges and reimbursements for sample members during the demonstration. Interviews with program administrators suggested that receipt of additional reimbursements for services to sample members was highly unlikely. No baseline data on charges and reimbursements were collected.

Statistical Methods

Both "unadjusted" and "adjusted" differences in mean values of measures of service use and reimbursement were calculated for the prepaid and traditional Medicaid groups. Unadjusted differences are differences in average values for the two groups when no statistical controls are applied for differences in baseline characteristics. The groups were similar at baseline with respect to observable characteristics; this similarity supports the use of unadjusted means in making comparisons with the entire sample. Adjusted differences are the differences that remain after a statistical relationship between the service use or reimbursement variable and the set of covariates listed in Table 2 has been estimated. Adjusted comparisons control for remaining baseline differences in beneficiary characteristics. The use of adjusted differences reduces the variance in the estimate of the difference between the experience of beneficiaries in prepaid plans and that of beneficiaries in traditional Medicaid. When estimating adjusted differences in the use and nonuse of community-based programs, we employed logit analysis and calculated differences by setting independent variables equal to their mean values. When estimating differences in number of visits, we used multiple regression analysis in which the dependent variable was the number of visits reported at the follow-up interview minus the number reported at baseline. By using the change in number of visits to a community mental health program as the dependent variable, we avoided estimation problems that can occur when the distribution of the dependent variable is characterized by a high proportion of zero values, as is the case for visits to some types of treatment programs. The coefficient on the prepaid vs traditional Medicaid variable was interpreted as the difference in the change in the measure for the two groups.

Multiple regression techniques were also used to test hypotheses relating to HMO write-offs. However, because no baseline data on reimbursements were collected, the adjusted difference relates to the average write-off during the period covered, rather than to the change from a baseline to a follow-up value. Because this analysis relates only to beneficiaries with some service use, and there is no concentration of dependent variable values at zero, impacts are not measured in terms of changes from baseline measures.

A drawback to the use of adjusted differences (whether calculated by logit or multiple regression analysis) is that fewer observations are available for analysis because of the cumulative impact of missing values for variables. To minimize the loss of data, mean sample values for independent variables were substituted for missing values; this was done for five or fewer observations for each variable. Where data were missing for dependent variables, observations were dropped from the analysis. Consequently, the sample sizes used in estimating adjusted effects vary; sample sizes are reported in the tables, along with significance levels.

Results

Use of Community-Based Mental Health Treatment Programs

Use of community-based treatment programs was measured on the basis of interview responses at baseline and follow-up; each measure covered the 3-month period prior to the interview. At the baseline interview, a significantly larger percentage of the traditional Medicaid group than of the prepaid group reported visiting a day treatment center (12.90% vs 7.07%), and the difference in average number of visits

approached statistical significance as well (3.47 vs 2.07, respectively). The groups were similar at baseline for other measures of use (Table 3).

A comparison of the unadjusted means for percentage of beneficiaries reporting a visit and average number of visits during the demonstration (Table 3, columns 4 and 5) indicates that prepaid group members used the services less often, but none of these differences were statistically significant (column 6). Two differences approached statistical significance: beneficiaries in prepaid plans reported fewer visits, on average, to community mental health centers (.29 vs .53) and a smaller percentage reported visits to crisis centers (4.18% vs 7.10%).

The number of visits generally declined over the course of the demonstration for the prepaid group, except for drop-in centers, where self-reported use increased (Table 3, columns 7 and 8). There was a reduction of .26 in the average number of community mental health center visits by the prepaid group, compared with an increase of .11 for the traditional Medicaid group; the difference in these changes reached statistical significance at the .07 level (column 9). Changes in the use of day treatment were significantly different at the .08 level, owing to a large decline in the number of visits by the traditional Medicaid group.

Column 10 of Table 3 shows estimates of the difference between the two groups in the probability of using each treatment program during the demonstration, after baseline characteristics of the respondents were controlled for. None of the estimated differences approach statistical significance (column 12). Column 11 presents estimates of differences (prepaid group vs traditional Medicaid group) in the changes in number of visits from baseline to follow-up period when respondent characteristics were controlled for. Again, these estimated differences are not statistically significant.

In summary, these results suggest that the enrollment of Medicaid beneficiaries in prepaid plans had a limited impact on the beneficiaries' propensity to use community-based mental health treatment programs and very little impact on the average number of visits to these programs. A consistent, statistically significant pattern of less use by prepaid group members was not evident when we examined changes from baseline values or when we controlled for the baseline characteristics of respondents in estimating differences.

TABLE 2—Covariates Used in the Estimation of Regression-Adjusted Differences (Measured at Baseline)

	Mean	SD
Prepaid vs traditional Medicaid (proportion prepaid)[a]	.50	.50
Age	41.56	11.61
Sex (proportion female)[a]	.56	.50
Race (proportion White)[a]	.84	.37
Education (highest grade)	11.94	2.48
Living arrangement (proportion living alone)[a]	.38	.49
Marital status (proportion married)[a]	.05	.21
Income (dollars per month)	397.18	255.81
Participation in Medicare (proportion Medicare)	.52	.50
Possession of private insurance (proportion private insurance)	.11	.31
Number of mental health admissions (prior year)	.47	1.23
Number of physical health admissions (prior year)	.41	1.05
Number of chemical dependency admissions (prior year)	.05	.22
Mental health outpatient visits (past 3 months)[a]	4.96	7.42
Physical health outpatient visits (past 3 months)	3.03	5.44
Chemical dependency outpatient visits (past three months)	1.11	5.30
General health status index (range, 4–16)[a]	10.19	2.36
Number of comorbidities	3.59	2.59
Physical functioning index (range, 0–9)	4.43	3.04
Global assessment scale score[a]	55.47	15.21
Community functioning (proportion arrested, jailed, assaulted, or who spent nights on the street)	.22	.41
Attempted suicide (proportion)	.07	.26
Number of days in plan[a]	328.70	52.69
Scores on SADS A–F subscales[a]		

Note. A covariate was also entered that interacted with mental and physical health care admissions and visits.
[a]These baseline characteristics were controlled for in the regression analysis.

Charges and Write-offs for Community-Based Mental Health Treatment Programs

Drop-in and crisis centers were omitted from this analysis because billing records did not exist, were incomplete, or were of questionable reliability. The data on the remaining two programs cover the entire period in which the sample member was a participant in the demonstration, rather than the 3-month recall period used in the analysis of data on service use. The analysis has several limitations. Because comparisons of mean values are reported only for users of services whose records were accessible, the sample sizes are small. Also, the protection provided by randomization does not apply to these comparisons (columns 3 and 4) because they rely only on self-selected samples of individuals who reported some service utilization. Therefore, it is possible that observed differences (column 5) could reflect respondent characteristics as well as demonstration impacts. To address this possibility, we also estimated differences using regression analysis (column 7) to control for selected baseline characteristics (indicated in Table 2), where sample size permitted.

As Table 4 indicates (under the heading "Total"), charges per person were higher, but not significantly higher, for prepaid enrollees than for traditional Medicaid beneficiaries. However, the proportion of the bill not paid (the write-off) was significantly larger for prepaid enrollees (.58) than for traditional Medicaid beneficiaries (.30). As a result, the average reimbursement received for prepaid enrollees was less. The estimated difference in write-off remains approximately the same after adjustments are made for respondent characteristics, and is still significant.

To investigate whether HMOs were hesitant to reimburse for particular treatment modalities or types of providers, we analyzed the data after disaggregation by these variables. Although the sample sizes after disaggregation are small for some categories, all results are reported to facilitate comparisons. When data were disaggregated by type of treatment, prepaid enrollees had higher charges and higher write-offs for psychotherapy, drug therapy, and day treatment. Despite small sample sizes, the difference in write-offs is significant at the .01 level for drug therapy and at the .06 level for psychotherapy after

TABLE 3—Use of Community-Based Mental Health Treatment Programs at Baseline vs Demonstration

	Mean Values at Baseline			Mean Values at Follow-up			Mean Value of Change in Number of Visits from Baseline to Follow-up			Controlling for Sample Characteristics		
	(1) PP	(2) TM	(3) *P* Value	(4) PP	(5) TM	(6) *P* Value	(7) PP	(8) TM	(9) *P* Value	(10) Difference in Percentage with Visit at Follow-up (PP–TM)	(11) Difference in Mean Value of Change in Number of Visits from Baseline to Follow-up (PP–TM)	(12) *P* Value
Percentage with visits in previous 3 months												
Community mental health center	9.00	7.79	.69	9.03	9.35	.99	...	...	...	.00	...	.99
Crisis center	4.82	7.10	.30	4.18	7.10	.16	...	...	...	2.25	...	.47
Drop-in center	4.84	3.23	.41	5.47	5.81	.99	...	...	...	−.78	...	.78
Day treatment center	7.07	12.90	.02	6.43	7.74	.63	...	...	...	.91	...	.64
Number of visits in previous 3 months												
Community mental health center	.55	.42	.52	.29	.53	.13	−.26	.11	.07	...	.03	.92
Crisis center	.09	.11	.67	.07	.10	.35	−.02	−.01	.79	...	−.10	.15
Drop-in center	.26	.62	.33	.46	.48	.92	.20	−.14	.51	...	.39	.54
Day treatment center	2.07	3.47	.11	2.00	1.98	.99	−.07	−1.49	.08	...	1.08	.79

Note. There are 311 individuals with complete baseline and follow-up data for the prepaid (PP) group and 310 for the traditional Medicaid (TM) group. In columns 3, 6, 9, and 12 the *P* values refer to significance levels for PP–TM differences.

adjustments are made for differences in respondent characteristics.

When data were disaggregated by provider type, write-offs were consistently higher for prepaid enrollees across all provider types; significant differences occurred for physician and "other clinician" write-offs. For these provider types, charges were also higher for prepaid enrollees, but not significantly so. There were enough observations in these two categories to permit estimation of regression-adjusted means as well. The adjusted differences were comparable to the unadjusted differences; the difference for physician write-offs remained statistically significant.

In summary, community-based mental health treatment programs had higher write-offs for Medicaid beneficiaries enrolled in prepaid plans. When data were disaggregated by provider, significant write-offs occurred for treatment by physicians. One possible explanation for this pattern of higher write-offs is that community-based treatment programs attempted to bill prepaid plans at higher rates than are allowed for traditional Medicaid beneficiaries. Charges for prepaid beneficiaries are higher, but the differences in charges are not statistically significant, whereas differences in write-offs are significant. It is important to reiterate, however, that small sample sizes, together with the large variance that exists in average charges, reduce the likelihood that statistically significant differences in the charge data will be observed.

Discussion

With respect to our first general hypothesis, we found no strong evidence that Medicaid beneficiaries with severe mental illness who were enrolled in prepaid health plans used community-based mental health treatment programs any differently than did other Medicaid beneficiaries. There are a number of possible explanations for this finding. First, beneficiaries were enrolled in prepaid plans for less than a year, and this may have limited the ability of the plans to alter the utilization behavior of enrolled beneficiaries. Our interviews with demonstration administrative staff and the staff of the prepaid plans indicate that plan enrollees continued to receive treatment from their predemonstration providers until the end of an illness episode or until transfer to a plan provider was arranged by mutual agreement. Thus at least some beneficiaries were able to continue as clients of community-based programs with little or no interruption even after their enrollment in prepaid plans.

A second explanation relates to the networks of providers available under the different prepaid plans and the freedom that beneficiaries had to choose among the available plans. As noted previously, the Hennepin County plan provided access to most community-based treatment programs, and Blue Cross/Blue Shield's provider network also included many community-based programs. Furthermore, the plans were not allowed to require that enrollees seek a physician referral to access the services of these programs.

The evidence supports the second general hypothesis: write-offs (based on unadjusted means) were significantly higher for prepaid plan enrollees than for other Medicaid beneficiaries across virtually all categories of provider and service. At least some of these differences remained significant after individual characteristics were controlled for, where sample size permitted. However, this conclusion must be interpreted with caution because there may be differences in the baseline characteristics of sample members that were not adequately controlled for in the regression analysis. Also, there is some evidence suggesting that charges were higher for prepaid plan enrollees, as well.

There is a plausible scenario that would explain our finding of no difference between the two groups in the use of community-based programs, along with higher write-offs for prepaid beneficiaries. Program policies encouraged the prepaid plans not to disrupt treatment that was ongoing at the time the demonstration was

TABLE 4—Charges and Write-offs by Treatment and Provider Type during Demonstration

	Sample Size		Comparison of Unadjusted Means				Comparison of Means when Sample Characteristics, were Controlled for	
			Mean Values					
	(1) PP	(2) TM	(3) PP	(4) TM	(5) Difference	(6) *P* Value	(7) Difference	(8) *P* Value
Total	61	57						
Charges, $			2572.06	2228.45	343.61	.59	660.74	.31
Write-off			.58	.30	.28	.00	.24	.00
Disaggregated by treatment								
Psychotherapy	18	19						
Charges, $			927.43	525.91	401.52	.11	220.07	.52
Write-off			.68	.33	.35	.01	.35	.06
Group Therapy	6	5						
Charges, $			336.46	895.70	−559.24	.01	NA	NA
Write-off			.49	.74	−.25	.24	NA	NA
Drug Therapy	35	28						
Charges, $			241.75	197.73	44.02	.36	−42.91	.41
Write-off			.79	.40	.39	.00	.46	.00
Diagnostic	10	8						
Charges, $			173.05	237.88	−64.83	.37	NA	NA
Write-off			.86	.75	.11	.49	NA	NA
Day Treatment	19	23						
Charges, $			5582.87	3759.71	1823.16	.14	761.41	.66
Write-off			.37	.15	.22	.02	.11	.38
Disaggregated by provider								
Physician	36	31						
Charges, $			453.01	282.60	170.41	.25	1.22	.61
Write-off			.80	.41	.39	.00	.43	.00
Licensed clinical psychologist	8	9						
Charges, $			258.56	340.83	−82.27	.54	NA	NA
Write-off			.74	.43	.31	.11	NA	NA
Master of Social Work	12	13						
Charges, $			640.21	575.23	64.98	.80	NA	NA
Write-off			.70	.62	.08	.58	NA	NA
Registered Nurse	5	3						
Charges, $			462.50	593.75	−131.25	.74	NA	NA
Write-off			.80	.63	.17	.56	NA	NA
Other clinicians	29	31						
Charges, $			4422.81	3410.62	1012.19	.33	1166.88	.39
Write-off			.32	.14	.18	.03	.13	.18

Note. PP = prepaid group; TM = traditional Medicaid group. Write-off is defined as (charges minus reimbursement)/charges.

initiated. Policies also did not allow prepaid plans to require prior authorization by a physician for referral to community-based mental health treatment programs; such a requirement is a typical utilization management technique employed by HMOs. The fact that they were not allowed to require such authorization may have caused the prepaid plans to rely heavily on retrospective denial of payment for services rendered by community-based programs as a means of controlling expenditures. Some programs may have accepted higher write-offs for prepaid enrollees because the subsidy they received from the public sector was based in part on the volume of services they provided. In fact, two prepaid plans attempted unsuccessfully to negotiate reimbursement rates with community-based programs, leading these plans to conclude that the programs had little incentive to aggressively pursue private reimbursement for the services they provided.[13] □

Acknowledgments

Financial support for this research was provided by Hennepin County (Minnesota), the National Institute of Mental Health, the Bush Foundation, the state of Minnesota, and the Hennepin Faculty Associates Young Investigators Program. Dr Lurie is a Henry J. Kaiser Family Foundation Faculty Scholar in general internal medicine.

The authors wish to thank the administrators of community-based treatment programs in Hennepin County for providing access to financial data; the staff of Hennepin County and of the Medicaid program in the State of Minnesota for facilitating access to Medicaid records and for providing their support for the research; and the administrators of the prepaid plans and the Medicaid beneficiaries in our sample for their willingness to participate in in-person interviews. We also acknowledge the excellent assistance in data collection provided by Colleen King and Susan Ricketts.

References

1. Schlesinger M. On the limits of expanding health care reform: chronic care in prepaid settings. *Milbank Q.* 1986;64:189–215.
2. Flinn D, McMahon T, Collins M. Health maintenance organizations and their implications for psychiatry. *Hosp Community Psychiatry.* 1987;38:255–263.
3. Fink P, Dubin W. No free lunch: limitations on psychiatric care in HMOs. *Hosp Community Psychiatry.* 1991;42:363–365.
4. Lehman A. Capitation payment and mental health care: a review of the opportunities and risks. *Hosp Community Psychiatry.* 1987;38:31–38.
5. Meier G. HMO experiences with mental health services to the long-term emotionally disabled. *Inquiry.* 1981;18:125–138.
6. Sabin J. Research findings on chronic mental illness: a model for continuing care in health maintenance organizations. *Compr*

Psychiatry. 1978;19:83–94.
7. Shadle M, Christianson J. The organization of mental health care delivery in HMOs. *Administration Men Health*. 1988;15:201–225.
8. Bittker T. Key concepts in prepaid mental health service delivery. *HMO Mental Health Newsletter*. 1987;2:1–5.
9. Raynes A. Contracts and capitation: the pros and cons of external linkages. Presented at HMO Mental Health: A Course for Administrators and Clinicians, Harvard Medical School, Department of Continuing Education, Boston, Mass, October 26–28, 1988.
10. Moscovice I, Finch M, Lurie N. Minnesota: plan choice by the mentally ill in medicaid prepaid plans. In: Scheffler R, Rossiter L, eds. *Advances in Health Economics and Health Services Research: Risk-Based Payments Under Public Programs*. Greenwich, Conn: JAI Press; 1989;10:265–280.
11. Medicaid Demonstration Design and Development Task Force. *A Plan for a Medicaid Demonstration Project in Hennepin County*. A report to the Hennepin County Board of Commissioners, Minneapolis, Minn; 1985.
12. Christianson J, Lurie N, Finch M, Moscovice I. Mandatory enrollment of medicaid-eligible mentally ill persons on prepaid plans: the Minnesota demonstration project. *Administration and Policy in Mental Health*. 1988;16:51–64.
13. *The Medicaid Demonstration Project in Hennepin County, Minnesota: Year One*. A report to the Hennepin County Board of Commissioners, Minneapolis, Minn; 1987.

Commentary: Caring for the Indigent Mentally Ill—New Strategies and Old Problems

ABSTRACT

Christianson and his colleagues examine how Medicaid beneficiaries receive mental health services in HMOs by analyzing two important aspects of service delivery: the use of community-based treatment programs by Medicaid beneficiaries enrolled in health maintenance organizations (HMOs) and the reimbursement levels paid to these programs by HMOs. The hypotheses studied are complex issues that concern mental-health advocates and providers. Traditional community-based mental health services have always struggled to maintain their presence in the health care field, having to contend with changing funding priorities and more serious and multiple problems presented by their patients. For prepaid plans to work effectively for the indigent mentally ill, the complex issues have to be made clear and acknowledged as meaningful variables. (*Am J Public Health*. 1992;82:796–798)

Ernesto Ferran, Jr, MD

Dr Christianson and his colleagues analyze two important aspects of mental health care delivery for those who have serious mental disorders and are indigent.[1] They want to know whether Medicaid beneficiaries, once enrolled in prepaid medical plans, will utilize mental health services to the same degree as do non-enrolled Medicaid beneficiaries, and whether the plans will pay the mental health providers the amounts requested of them. Advocates for the indigent mentally ill population would define these questions as questions of access to mental health services and protection of funding for providers traditionally vulnerable to new financing schemes. The issues are not only complex, but also mysterious, and they lend themselves to contentiousness and partisanship.

What I want to address are the factors that make these issues complex: the fact that psychiatric standards of care are difficult to quantify; that mental health care providers, in the absence of conclusive outcome studies for traditional services, are promoting new strategies for providing services to the seriously mentally ill; and that agreement on reimbursement for these strategies will continue to be elusive.

By addressing their first question without a directional hypothesis, the authors avoid being drawn into the argument that relates utilization of outpatient mental health care to outcome. It makes sense that they do so: our fascination with health care expenditures as an object of study or control becomes problematic for disorders of behavior, emotion, or thought, and the relevant standards of care continue to be poorly understood. Consider the near decade-long effort to create adequate diagnosis-related groups for psychiatric disorders. Witness the relatively recent controversies surrounding for-profit psychiatric hospitals, and the accompanying arguments on appropriate inpatient care vs expected or egregious pricing strategies. Have you not wondered, at some point, what it is that the most vulnerable segment of America's homeless really needs—housing, employment, or long-term psychiatric inpatient care? In other words, the quest to quantify robust standards of care for psychiatric disorders continues to generate more controversy than consensus.

The second hypothesis almost speaks for itself: Write-offs will be higher because prepaid plans will not pay the full costs of care. Why wouldn't they? Why even consider that they wouldn't? If outcome data for outpatient mental health services are so scant, and if "managed-care" pricing strategies (e.g., using diagnostic categories to determine allowable units of service) are de-

Ernesto Ferran, Jr, MD, is with Gouverneur Hospital, New York, and with New York University Medical Center.

Requests for reprints should be sent to Ernesto Ferran, Jr, MD, Director of Psychiatry, Gouverneur Hospital, 227 Madison Street, New York, NY 10002.

Weight Gain Prevention and Smoking Cessation: Cautionary Findings

ABSTRACT

Objectives. Weight gain is a consistent sequela of smoking cessation. A successful intervention might attract smokers who fear weight gain. If the gain causes smoking relapse, such an intervention might reduce smoking relapse risk.

Methods. Using a sample of 158 smokers who completed a 2-week smoking treatment program, we compared an innovative weight gain prevention intervention with both a nonspecific treatment and standard treatment. Subjects were assessed on weight and smoking behavior and followed for 1 year.

Results. A disturbing, unexpected finding was that subjects in both the innovative and nonspecific conditions had a higher risk of smoking relapse than did standard treatment subjects. Some differences were observed between abstinent and smoking subjects in weight gain by treatment condition.

Conclusions. Both active interventions may have been so complicated that they detracted from nonsmoking. Also, caloric restriction may increase the reinforcing value of nicotine, a psychoactive drug, thereby increasing smoking relapse risk. The magnitude of weight gain after smoking cessation may not merit interventions that increase smoking risk. Perhaps attitudinal modifications are the most appropriate. (*Am J Public Health.* 1992;82: 799–803)

Sharon M. Hall, PhD, Chrystal D. Tunstall, PhD, Katharine L. Vila, MSW, and Joanne Duffy, PhD

Introduction

Weight gain is a consistent sequela of smoking cessation. Of 43 studies of weight change after smoking initiation or cessation, 37 (86%) report that initiation of smoking reduces weight and cessation increases it. The cause of the relationship between smoking cessation and weight gain is unknown although changes in both metabolism and food intake are implicated. Smokers, especially young women who smoke, are aware of nicotine's weight-suppressing effects.[1] Nevertheless, studies of the effects of weight gain on relapse sometimes report surprising findings. For example, investigators[2–4] have found that weight gain during early abstinence predicted long-term abstinence, not relapse.

A successful intervention for weight gain prevention might attract smokers who fear weight gain and could potentially reduce relapse. In those few cases in which weight gain would be sufficient to pose a medical problem, such an intervention would prevent increased risk.

We designed an innovative intervention based on a model that suggests that adherence to behavioral prescriptions is facilitated by simplicity and by a close fit between self-management techniques and the problem behavior.[5] In this study, we report a comparison of this intervention with nonspecific and standard treatment conditions. The nonspecific condition was equiva ent to the innovative condition in time and therapeutic attention. The standard treatment control condition was designed to represent standard care of cessation-induced weight gain.

We proposed the following hypotheses: (1) Abstinent smokers in the innovative intervention would gain less weight than those in the nonspecific intervention, who would gain less weight than abstinent smokers in the standard treatment control condition. (2) Abstinence rates in the innovative intervention would be greater than those in the nonspecific intervention or in standard treatment because the close fit of the therapeutic techniques to cessation-induced gain in the innovative condition would be less likely to interfere with abstinence.

We also examined activity levels and, in a subsample, changes in food intake to determine whether subjects adhered to the instructions in the interventions.

Methods

Subjects and Personnel

Subjects were 49 male and 131 female smokers, all of whom met at least one of two risk factors for weight gain: they had to have smoked at least 10 cigarettes per day in the past week or been at least 10% above ideal weight at their maximum.[2] We excluded potential subjects with cardiovascular or pulmonary disease, diabetes, hospitalization for major mental illness within the last year, or any chronic condition that would influence eating or activity levels. Of these 180 subjects, 165 (92%) completed the smoking cessation treatment and were randomized to one of the weight gain prevention interventions.

The authors are with the University of California and the Veterans Affairs Medical Center in San Francisco.

Requests for reprints should be sent to Sharon M. Hall, PhD (116A), Veterans Affairs Medical Center, 4150 Clement Street, San Francisco, CA 94121.

This paper was submitted to the Journal June 10, 1991, and accepted with revisions November 12, 1991.

There were four female therapists. Two, a master's-level social worker and a doctoral-level psychologist, treated 159 (88%) of the subjects. The other two therapists were a master's-level health educator and a master's-level psychologist. Before beginning the intervention, therapists led at least one training group and observed other therapists' groups. During the intervention, therapists followed a detailed procedures manual. Two female registered dietitians and a male exercise consultant also participated.

Assessments

Assessments were held at pretreatment, at week 2 (the end of smoking treatment), at week 6 (the end of the weight gain prevention intervention), and at weeks 12, 26, and 52 from study start. At each assessment, subjects reported the average number of cigarettes smoked per day in the past week and provided expired air carbon monoxide samples to verify abstinence. At pretreatment and at the week 52 assessment, blood cotinine levels were determined as well. During each assessment week, subjects monitored their activity levels by recording time spent in bouts of activity on a schedule of 49 activities developed from McArdle, Katch, and Katch.[6] Body weights were taken at all assessments on the same balance beam scale. Twenty-four-hour dietary recalls, later processed for nutrient analyses by Nutrition and Diet Services,[7] were obtained from the last 91 subjects to enter the study. At baseline, subjects indicated the number of times they had attempted to quit smoking, their weight gain during past attempts, and whether they were concerned about and expected weight gain during the current attempt.

Procedure

Subjects were randomly assigned, within their time constraints, to groups of five to eight participants. During weeks 1 to 2, all groups received a smoking treatment program of seven 1½-hour sessions that combined aversive smoking and relapse prevention skill training. Until the last smoking treatment session, both therapist and group members were blind to which weight gain prevention intervention the group was assigned.

Subjects assigned to the innovative and nonspecific interventions participated in five weight gain prevention sessions held during weeks 3 to 6. During these sessions, discussions of purely smoking-related phenomena were minimized to focus on weight gain prevention and to keep about equal the amount of time spent on smoking cessation in the three conditions. Other than the assessment at week 6, no meetings were held in the standard treatment condition during these 4 weeks.

Innovative Intervention. There were three components to this intervention.

1. Daily monitoring of weight and contingent caloric reduction. Before the first session of the intervention, the dietitian computed for each subject a calorie level (the "food plan") designed to result in a weight loss of 2 lb per week based on body weight, age, sex, and activity level. The dietitian discussed healthy food choices and provided tips on low-calorie, sweet-tasting foods and on low-fat alternatives to common foods. Subjects weighed themselves each morning without clothing. If their weight was within 2 lb of baseline weight, they ate whatever they liked; if they gained 2 lb or more, they were to use their food plan until baseline weight was achieved.

2. Individualized exercise plan. The exercise consultant explained the principles of aerobic exercise and the benefits of fitness, and he helped subjects select an activity. Subjects were encouraged to exercise three or more times a week at an aerobic level.

3. Behavioral self-management principles. The subjects discussed triggers that resulted in uncontrolled eating. If appropriate, new responses were role-played. Subjects were asked to limit eating to one place at home and at work, and to slow eating by taking a single bite at a time.

Nonspecific Control. The nonspecific control condition included the elements of a credible weight gain prevention intervention without the specific elements of the innovative condition. The nonspecific elements were (1) a rationale based on gaining insight into eating styles, (2) a structured program based on insight-oriented discussion, (3) nutritional and exercise information, (4) group support, and (5) therapeutic attention. The sessions centered around both self-tests adapted from the Smokers' Self-test Kit[8] and informational presentations by the nutritionists and the exercise consultant. The dietitian presented general nutritional information about calories and good nutrition, noting that caloric reduction was necessary for weight loss. The fitness consultant presented fitness principles and the health benefits of exercise. However, there was no attempt to individualize the information. The therapist did not give specific behavioral prescriptions. Rather, she encouraged subjects to decide how they would use feedback provided by the group, and she referred questions about specific techniques back to the group. The number and spacing of sessions were the same as those in the innovative condition.

Standard Treatment Control. Subjects received an information packet on good nutrition and exercise at the last smoking treatment session. This material was not targeted for smoking cessation–induced weight gain.

Results

Preliminary Analyses

As previously noted, 165 of the 180 subjects who began treatment completed the smoking cessation portion. Dropouts did not differ from the rest of the sample in age, gender, ethnicity, level of education, socioeconomic status, number of previous quit attempts, cessation-induced weight gain, pretreatment carbon monoxide, cigarette intake, or body weight. The data from one subject of each pair of six close relatives were randomly selected to use in the data analyses. Data from an additional subject who was later diagnosed as bulimic were discarded. Resultant sample size was 158. Sizes by intervention condition were innovative intervention, 53; nonspecific control, 51; and standard treatment control, 54. When asked if they expected to gain weight after quitting, 112 of these 158 subjects (71%) said yes. One hundred thirty-eight (87%) reported at least one earlier quit attempt, 86 of whom (62%) reported a weight gain that averaged 6.35 kg (SD = 3.86 kg). Almost all the subjects (92%) were concerned about gaining weight when they quit.

Treatment conditions were compared on the same demographic, smoking, and weight history variables used to compare treatment completors and dropouts, as well as on postcessation body weight and postcessation cigarette smoking rates (see Table 1). Only two differences emerged. First, conditions differed on mean number of cigarettes reported smoked per day during the pretreatment week ($F[2,155] = 5.06$, $P < .01$). Subjects in standard treatment smoked significantly more cigarettes than did those in the other two conditions. Second, the three conditions had significantly different variances on pretreatment body weight ($P < .01$). Subjects in the innovative condition had a significantly greater variance in both these measures than did subjects in the nonspecific condition. In tests of the hypotheses, these variables—pretreat-

ment smoking rate and pretreatment body weight—were used as covariates in preliminary analyses if they correlated with the dependent variable, and they were dropped if they did not explain significant proportions of variance or have significant interactions with independent variables. Due to a lack of significant main effects, or interactions, these variables were ultimately dropped from all analyses, except those of abstinence status. No differences due to therapists were found when the main effects of therapist on smoking rate or body weight were examined.

Abstinence Status

At each assessment, subjects were coded as abstinent if they reported no cigarettes smoked during the prior week and had a carbon monoxide level below 10.5 ppm at the assessment. At week 52, subjects were coded as abstinent only if their blood cotinine levels were also less than 50 ng/mL. Chi-squares of abstinence status by treatment condition did not reach significance at any assessment (Table 2).

However, treatment conditions differed at baseline on number of cigarettes smoked, and baseline number of cigarettes did correlate with abstinence status at the followup assessments. The prototypical analysis was a multiple logistic regression with gender and number of cigarettes smoked at pretreatment as covariates. The dependent variable was abstinence status at each assessment week. Subjects with missing data were coded as smoking. These analyses indicated that, at weeks 6, 12, 26, and 52, treatment condition contributed significantly to the abstinence rate. Subjects in the two active interventions had a greater risk of smoking than did subjects in the control condition. Effects for the two active treatment conditions did not differ from one another.

The odds ratios, associated confidence intervals (CIs), and probability levels are shown in Table 3. The nonspecific condition differed significantly ($P \leq .05$) from the control condition at weeks 6, 12, and 52; the innovative condition differed significantly from the control at week 26. The combined active treatment conditions differed significantly from the control at each assessment.

Smoking more than 35 cigarettes per day at pretreatment significantly increased the risk of relapse at all assessments. The odds ratio ranged from 2.73 at week 6 (95% CI = 1.64, 4.57; $P < .001$) to 1.95 at week 52 (95% CI = 1.08, 3.52; $P < .03$). There were no significant effects for gender at any assessment.

TABLE 1—Demographic and Descriptive Data by Treatment Condition

Variable	Innovative Condition (n = 53)	Nonspecific Condition (n = 51)	Standard Treatment Condition (n = 54)
Age, y			
Mean	40.68	41.22	39.24
SD	8.71	8.96	8.92
Cigarettes per day			
Mean	25.65	24.91	31.59
SD	12.56	10.80	11.74
Pretreatment body weight, kg			
Mean	72.68	69.23	67.65
SD	17.72	12.38	12.51
Postcessation treatment body weight, kg			
Mean	73.82	69.98	67.97
SD	18.12	10.48	12.37
Gender			
Male	12 (23%)	17 (33%)	14 (26%)
Female	41 (77%)	34 (67%)	40 (74%)
Ethnicity			
Caucasian	48 (91%)	45 (88%)	50 (91%)
Other	5 (9%)	6 (12%)	4 (9%)

TABLE 2—Number and Percent Abstinent in Each Treatment Condition at Each Assessment

	Innovative Condition (n = 53)		Nonspecific Condition (n = 51)		Active Conditions Combined (n = 104)		Control Condition (n = 54)	
Assessment	n	%	n	%	n	%	n	%
Week 6	26	50	21*	41	47	45	31	57
Week 12	17	32	14*	25	31*	30	24	44
Week 26	11*	21	13	25	24	23	19	35
Week 52	11*	21	11	22	22*	21	19	35

*Comparison with control, $.05 < P < .10$.

Weight Change

The heterogeneity of variances in the weight data was not correctable by either logarithmic or square root transformations. We used a series of Kruskal-Wallis nonparametric analyses of variance on body weight change from week 2 to week 6 and to week 52 to determine if subjects differed by conditions and abstinence status. Kruskal-Wallis tests were also used to establish the lack of difference between conditions and smoking versus nonsmoking subjects on weight change from week 0 to week 2.

There were no differences among treatment conditions on weight change at either week 6 or week 52. Abstinent subjects differed from smoking subjects at both week 6 ($\chi^2[1] = 10.85$, $P < .01$; abstinent subjects, n = 78, mean = +0.79 kg, SD = 1.74; smoking subjects, n = 74, mean = +0.14 kg, SD = 1.32) and week 52 ($\chi^2[1] = 7.21$, $P < .01$; abstinent subjects, n = 31, mean = +2.59 kg, SD = 3.76; smoking subjects, n = 87, mean = +0.22 kg, SD = 3.42). Differences between abstinent and smoking subjects were not consistent across the three conditions, however. In the innovative condition, abstinent and smoking subjects did not differ at either week 6 (abstinent subjects, n = 26, mean = +0.08 kg, SD = 2.4; smoking subjects, n = 24, mean = 0.14 kg, SD = 1.26) or week 52 (abstinent subjects, n = 10, mean = +0.86 kg, SD = 3.95; smoking subjects, n = 35, mean = –0.31 kg, SD = 3.57). Differences between abstinent and smoking subjects were found in the nonspecific condition at both week 6 ($\chi^2[1] = 8.66$, $P < .01$; abstinent subjects, n = 21, mean = +1.20 kg, SD = 1.18; smoking subjects, n = 27, mean = +0.03

TABLE 3—Odds Ratios (ORs),[a] Confidence Intervals (CIs), and Probability Levels for the Two Active Conditions (Separately and Combined) vs the Standard Treatment

	Week 6				Week 12				Week 26				Week 52			
	OR	95%	CI	P	OR	95%	CI	P	OR	95%	CI	P	OR	95%	CI	P
Innovative vs Control	1.48	.89	2.45	.13	1.57	.95	2.60	.08	1.70	1.00	2.88	.05	1.59	.92	2.75	.10
Nonspecific vs Control	2.19	1.28	3.73	.01	2.05	1.19	3.52	.01	1.61	.93	2.79	.08	2.02	1.13	3.64	.02
Combined vs Control	1.76	1.12	2.77	.01	1.77	1.13	2.77	.03	1.66	1.04	2.63	.02	1.77	1.09	2.88	.02

[a]In the present study, the odds ratio is interpreted as the change in risk of relapse to smoking that is attributable to membership in the innovative or nonspecific condition, as compared with membership in the standard treatment control.

kg, SD = 1.39) and week 52 ($\chi^2[1] = 4.12$, $P < .04$; abstinent subjects, n = 7, mean = +3.35 kg, SD = 2.38; smoking subjects, n = 25, mean = +0.44 kg, SD = 3.60). The standard treatment control showed no differences at week 6 ($\chi^2[1] = 1.40$, $P < .29$; abstinent subjects, n = 31, mean = +1.12 kg, SD = 1.54; smoking subjects, n = 23, mean = +0.55 kg, SD = 1.27). However, differences in this condition approached significance at week 52 ($\chi^2[1] = 3.45$, $P < .06$; abstinent subjects, n = 14, mean = +3.61 kg, SD = 3.99; smoking subjects, n = 27, mean = +0.71 kg, SD = 3.09).

Calories and Activity Levels

Of the original subsample of 91 subjects on whom we attempted to collect nutritional data, 68 (75%) completed dietary recalls at every assessment and were therefore included in the repeated measures analyses of caloric intake. These subjects did not differ from either the whole sample or the sample of 91 on whom we attempted to collect data on cigarettes smoked or body weight.

An unacceptable level of heterogeneity of variances in baseline caloric intake was reduced to an acceptable level by a square root transformation. The major analysis was a repeated measures analysis of variance. The dependent variables were the square root transformation of caloric intake at baseline and at weeks 6, 12, 26, and 52. Independent variables were gender, treatment condition, and abstinence status at week 52.

A significant effect for treatment condition over time ($F[8,224] = 3.90$, $P < .001$) was found, as was a significant abstinence status × treatment condition × time interaction ($F[8,224] = 2.04$, $P < .05$). We found significant treatment effects at week 6 ($F[2,15] = 11.95$, $P < .01$) and week 26 ($F[2,15] = 2.92$, $P < .01$) but not at baseline, week 12, or week 52. At week 6, subjects in the innovative condition ate significantly fewer calories than those in the nonspecific condition ($t[42] = 2.48$, $P < .03$; innovative condition, mean = 1687 cal, SD = 707; nonspecific condition, mean = 1810 cal, SD = 604) and those in the standard treatment control ($t[42] = 4.29$, $P < .01$; mean = 2817 cal, SD = 976). Differences in caloric intake between the nonspecific treatment condition and the standard treatment control were not significant ($t[46] = 1.83$, $P < .10$). At week 26, none of the paired comparisons was significant.

For subjects smoking at week 52, the time × treatment condition interaction was significant ($F[8,164] = 2.77$, $P < .01$). There were also significant treatment condition × gender interactions at baseline ($F[2,41] = 3.65$, $P < .04$), explained primarily by significantly greater caloric intake by men in the innovative condition than by subjects of either gender in any of the five other gender-treatment condition groupings. Repeated measures analysis of variance indicated no significant effect for activity level.

Discussion

The higher smoking rates at follow-up in the two conditions that received an active treatment are troubling. The odds ratios for the risk of smoking for the two active treatment conditions combined were consistently and significantly greater than the odds ratios for the control condition at each assessment after week 2. The most likely explanation for this is that both interventions were sufficiently complicated so that subjects' attempts to apply them detracted from attempts to maintain abstinence. A second explanation is that the caloric reduction encouraged by both active conditions encouraged smoking. Caloric restriction may increase the reinforcing value of psychoactive drugs in laboratory animals.[9]

The failure to find a significant difference between abstinent and smoking subjects in the innovative condition suggests that this condition prevented weight gain, especially when this lack of difference is compared with the significant differences found between abstinent and smoking subjects in the nonspecific condition. Differences in the control condition, however, also failed to reach significance, rendering interpretation tentative. In the innovative condition there was evidence of early compliance with dietary instructions, especially among women and abstinent subjects. The interventions had no measurable affect on self-reported activity.

The limitations of this study are important. First, the results obtained apply to one program; other programs may be more successful. Second, the finding of increased relapse rates in the two active treatment conditions must be viewed with caution. The design of the study, in which differing combinations of behavioral therapists, nutritionists, and exercise therapists treated subjects in each condition, did not lend itself to a thorough analysis of intervenor × treatment condition interaction effects. These effects are known to influence the magnitude of differences among treatment conditions, at least in traditional analysis of variance models.[10] So far as we could find, nothing has been published about these effects in logistic models. Inspection of the data and those analyses we did perform did not suggest consistent intervenor effects, and significant main effects for intervenors were not found. Such analyses, however, cannot substitute for a complete factorial analysis of their differences.

Finally, it should be noted that the average gain for ex-smokers over 1 year is about 4 lb greater than that expected for continuing smokers. One study reported that about 3% of quitters gained 20 lb or more.[11] Such small weight gains do not appear to merit interventions that are extensive or entail risk. Perhaps attitude

changes about weight gain are more appropriate interventions than treatment of that gain. □

Acknowledgments

This study was funded in part by grants DA02356 and DA00065, both from the National Institute on Drug Abuse, and by a Research Career Scientist Award from the Department of Veterans Affairs to Dr. Hall.

The authors wish to thank Dia Hartz for her help in data analysis, and Patti Kuzimoto, Kathy McHenry, Afsaneh Shayestah, and Alan Tower for their contribution as consultants.

References

1. *The Health Consequences of Smoking. Nicotine Addiction: A Report of the Surgeon General.* Rockville, Md: US Department of Health and Human Services. 1988: 431–432. DHHS publication (CDC)88–8406.
2. Hall SM, Ginsberg D, Jones RT. Smoking cessation and weight gain. *J Consult Clin Psychol.* 1986;54:342–346.
3. Gritz ER, Carr CR, Marcus AC. Unaided smoking cessation: The Great American Smokeout and New Year's Day quitters. *J Psychosoc Oncology.* 1988;6:217–234.
4. Hughes JR, Gust SW, Skoog K, Keenan RM, Fenwick JW. Symptoms of tobacco withdrawal: A replication and extension. *Arch Gen Psychiatry.* 1991;48:52–59.
5. Hall SM. Self-management and therapeutic maintenance: Theory and research. In: Karoly P, Steffan J, eds. *Improving the Long-term Effects of Psychotherapy.* New York, NY: Gardner Press; 1980:263–300.
6. McArdle WD, Katch FE, Katch V. *Exercise Physiology.* Philadelphia, Pa: Lea & Febiger; 1981:486–493.
7. Dwyer J, Suitor CW. Caveat emptor: assessing needs, evaluating computer options. *J Am Diet Assoc.* 1984;84:302–312.
8. Horn D. *Smoker's Self-testing Kit.* Atlanta, Ga: Dept of Health, Education, and Welfare; 1972. DHEW publication (HSM)72–7506.
9. Carroll M, Meisch RA. Increased drug-reinforced behavior due to food deprivation. In: Thompson T, Dews PB, Barrett JE, eds. *Advances in Behavioral Pharmacology.* Orlando, Fla: Academic Press; 1984:48–88.
10. Cris-Christoph P, Mintz J. Implications of therapists' effects for the design and analysis of comparative studies of psychotherapies. *J Consult Clin Psychol.* 1991;59:20–25.
11. *The Health Benefits of Smoking Cessation: A Report of the Surgeon General.* Rockville, Md: US Department of Health and Human Services; 1990:476–483. DHHS publication (CDC)90–8416.

Call for Abstracts for First International Conference on Community Health Nursing Research

The International Scientific Review Committee invites interested nurses to submit abstracts of papers or posters for presentation at the First International Conference on Community Health Nursing Research, to be held September 27 to 29, 1993, in Edmonton, Alberta, Canada. Papers and posters that deal with the development, design, or implementation of research in community/public health nursing and primary health care are welcome. The deadline for abstracts is *September 15, 1992.*

For abstract submittal forms and further information on the conference, contact the International Conference on Community Health Nursing Research, Edmonton Board of Health, Suite 500, 10216 - 124 St, Edmonton, Alberta, Canada T5N 4A3; tel. (403) 482-1965, fax (403) 482-4194.

Call for Abstracts for the Sixth International Conference on Indoor Air Quality and Climate

INDOOR AIR '93, the principal international conference on indoor air quality, will be held July 4 to 8, 1993, in Helsinki, Finland. The conference will offer an overview of the latest research results in the field, and anyone who is seriously involved in technological development, building sciences, regulatory issues, or other businesses related to indoor air is encouraged to attend.

Topics to be covered include public health, biosciences, occupational health, environmental engineering, building design and technology, energy consumption, building operation, material sciences, air chemistry, and regulatory issues.

Abstracts must be received by *October 1, 1992.* For a complete list of topics, an abstract submittal form, and further information on the conference, contact INDOOR AIR '93, PO Box 87, SF-02150 Espoo, Finland; fax 358 0 451 3611.

Improving Referral Compliance after Public Cholesterol Screening

ABSTRACT

Background. Noncompliance with referral to a physician for retesting and diagnosis is a concern in public cholesterol screening.

Methods. Participants (n = 2109) were referred by a health professional or lay communicator and randomly assigned to a coupon offer, referral reminder letter, or control group. A questionnaire was completed at screening, and a telephone interview was conducted 5 months later.

Results. Physician visit rates showed no professional or lay differences. For "no history" subjects, the behavioral interventions were effective compared with controls (coupon = 60.7% and reminder = 57.7% vs control = 46.1%). With professional counseling, only the coupon was effective; for lay counseling, both coupon and reminder yielded higher visit rates. Adjusted for sociodemographics, heart disease risk factors, and health perceptions, the intervention effects remained (professional–coupon offer: odds ratio [OR] = 1.94, 95% confidence interval [CI] = 1.21, 3.09; professional–reminder letter: OR = 1.04, 95% CI = 0.67, 1.63; lay–coupon offer: OR = 2.52, 95% CI = 1.52, 4.18; and lay–reminder letter: OR = 3.10, 95% CI = 1.83, 5.22).

Conclusions. For unaware participants, lay counselors and referral follow-up efforts tailored to specific cholesterol risk groups are indicated. (*Am J Public Health.* 1992;82:804–809)

Lois A. Maiman, PhD, Nancy G. Hildreth, PhD, MPH, Christopher Cox, PhD, and Philip Greenland, MD

Introduction

The National Cholesterol Education Program (NCEP) recommends blood cholesterol testing for all adults and has attempted to inform the public about the significance of high blood cholesterol levels in heart disease.[1] This effort and the recent availability of portable rapid cholesterol analyzers has led to increased participation in public screening.

Unanswered questions about mass cholesterol screening programs led the 1984 Consensus Development Conference to call for further research on the utility of these programs in order to formulate guidelines for public cholesterol screening under quality conditions.[2] Primary issues of concern were (1) whether screening can effectively link participants with elevated cholesterol to medical care, (2) whether physicians treat referrals according to current guidelines, and (3) the reliability of portable analyzers over extended time periods in screening field-type conditions.[3] Previous studies indicate that only 29% to 57% of participants with elevations follow through on referrals from public cholesterol screenings[4–6] and that physicians retest 33% to 48% who seek care.[4,7] The lack of acceptable accuracy and standardization in cholesterol measurement has been noted.[3]

As one of three Model Systems for Blood Cholesterol Screening Projects supported by the National Heart, Lung, and Blood Institute (NHLBI), the Rochester Project addressed all of these concerns. The present report focuses on interventions to increase referral compliance. The results concerning the increased responsiveness of physicians and acceptable reliability of portable analyzers have been reported elsewhere.[8,9]

The study examines whether health professional counselors at screening and follow-up interventions affect compliance with referral for retesting and diagnosis. The contributions of sociodemographic characteristics, coronary heart disease risk factors, and health motivations are also assessed.

Methods

Setting and Design

Voluntary, free cholesterol screenings were conducted in supermarket parking lots in Rochester, NY. Table 1 summarizes the study design. Participants, 20 through 65 years of age and meeting age-specific criteria for moderate- or high-risk cholesterol levels, were referred to a physician for retesting and diagnosis. Ineligibility occurred if conditions capable of causing cholesterol elevations were reported (pregnancy or lactation, or diabe-

At the time of the study, the authors were all with the University of Rochester School of Medicine and Dentistry, Rochester, NY. Lois A. Maiman is currently with the Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Human Development, Bethesda, Md. Philip Greenland is currently with the Department of Community Health and Preventive Medicine, Northwestern University Medical School, Chicago.

Requests for reprints should be sent to Lois A. Maiman, PhD, Head, Section on Prevention Research, Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Human Development, NIH, Executive Plaza North, Room 640, Bethesda, MD 20892.

This paper was submitted to the Journal January 28, 1991, and accepted with revisions December 31, 1991.

Editor's Note. See related editorial on p 785 of this issue.

TABLE 1—Study Design: High Blood Cholesterol Referral Compliance

Group	Screening[a]		Referral Interventions[b]							
			Communicator[c]		Behavioral[d]					
	Total Cholesterol	Baseline Survey	Professional	Lay	Coupon Offer	Reminder Letter	Control	n	Postscreening Survey[e]	n
Health professional–coupon offer	X	X	X		X			377	X	355
Health professional–reminder letter	X	X	X			X		369	X	343
Health professional–control	X	X	X				X	377	X	348
Subtotal								1123		1046
Lay–coupon offer	X	X		X	X			328	X	312
Lay–reminder letter	X	X		X		X		329	X	312
Lay–control	X	X		X			X	329	X	305
Subtotal								986		929
Total								2109		1975

[a]Eligibility: 20–65 years, not pregnant or lactating or having diabetes, liver disease, or hypothyroidism.
[b]Eligibility: moderate- or high-risk cholesterol level.
[c]Self-allocation at screening.
[d]Random assignment, mailed 3 days after screening.
[e]Five months after screening.

tes, liver disease, or hypothyroidism). Individuals without a regular source of medical care received a list of primary care physicians accepting new patients. Informed consent and a medical release for physician verification of cholesterol treatment were obtained at screening.

Total serum cholesterol was determined using a finger-stick blood sampling, and analyses were done on the Eastman Kodak Ektachem DT-60. Following standard procedures for quality control recommended by the Laboratory Standardization Panel of NCEP,[10] the between-run coefficients of variation ranged from 1.9% to 4.8% per month; average bias compared with the Abell-Kendall Reference Method ranged from −0.2% to 2.0%. Both precision and accuracy met recommended standards for cholesterol testing in the United States.[9]

At the initial level of the 2 × 3 factorial design, each participant received counseling from the first available health professional (i.e., nursing background) or lay communicator. Participants were unaware of the counselor's background. After each screening shift, referred participants were randomly assigned to a coupon offer, reminder letter, or control group.

Participants completed a self-administered questionnaire prior to receiving their test result and were interviewed by telephone approximately 5 months after the screening. If telephone contact was not made, a questionnaire was mailed. From June through October 1987, 7824 individuals were screened and 2109 referred. The follow-up survey involved 1975 subjects, and 1948 with complete data (92%) were included in the analyses. The cohort of referred participants unaware of elevated levels prior to screening (n = 1187) is the primary focus of this paper. Since public screening attracts individuals with known elevations, comparative findings for this subgroup (n = 761) are presented without tabular data in the interest of brevity.

To limit subject and communicator bias, the consent form described the survey as determining experiences after screening, the questionnaire and interview inquired about general health information as well as cholesterol-specific items and physician visit, and the communicators were "blind" to the interventions. Finally, interviewers were "blind" to a subject's intervention and screening risk group.

Measures

Predictor measures. The baseline questionnaire obtained data on sociodemographics, personal medical history of high blood cholesterol and other coronary heart disease risk factors, intention to visit a doctor if the cholesterol test was high, and perceptions of health and medical care. Screening cholesterol risk level for referral was defined according to age by blood concentration guidelines recommended at the time of the program (moderate risk: 20 through 29 years = 200 to 220 mg/dL; 30 through 39 = 220 to 240 mg/dL; 40 or over = 240 to 260 mg/dL; high risk: 20 through 29 = >220 mg/dL; 30 through 39 = >240 mg/dL; 40 or over = >260 mg/dL).[2]

Experience of counselor. Since messages for those with and without elevations were s[illegible]n content, except for the recommendation to consult a physician, experience of counselor was based on the total number of interactions. A categorical measure of counseling experience was defined for each referred subject by placement in his or her counselor's experience level across all participants counseled (1 = 1 through 74; 2 = 75 through 199; 3 = 200 through 299; 4 = 300 through 499; and 5 = ≥500).

Outcome measure. The follow-up interview asked whether participants had contacted their doctor about the screening test. A dichotomous dependent variable reported the proportion in each intervention group actually visiting their physician.

Self-reports and medical record data were compared for participants reporting a physician visit and identifying the source of care (n = 1082; 94%); 337 of 367 physicians returned record forms for 1001 subjects (completion rates of 92% and 93%, respectively). Physician documentation corroborated with 94% of the self-reports.

Interventions

Monitoring procedures ensured the correct-for-risk-level counseling and the accurate implementation of the follow-up interventions.

1. "Exit counseling communicators," half professional and half lay (n = 22), received the same length and content of training and served throughout the project. Counseling (3 to 5 minutes) included the following: cholesterol level and interpretation as "usual," "moder-

TABLE 2—Selected Sociodemographic Characteristics and Coronary Heart Disease Risk Factors of Study Participants by History of High Blood Cholesterol

Characteristics	Total Sample[a] (n = 1948)		No History (n = 1187)		Prior History (n = 761)	
	n	%	n	%	n	%
Age, y[b]						
20–44	473	24.3	359	30.2	114	15.0
45–54	427	21.9	270	22.8	157	20.6
55–65	1048	53.8	558	47.0	490	64.4
Sex						
Male	756	38.8	457	38.5	299	39.3
Female	1192	61.2	730	61.5	462	60.7
Education[c]						
<High school	159	8.2	87	7.4	72	9.5
High school	688	35.4	421	35.6	267	35.3
Some college	527	27.2	308	26.0	219	28.9
≥College graduate	567	29.2	368	31.1	199	26.3
Family income						
<$15 000	258	14.1	157	14.0	101	14.1
$15 000–$24 999	367	19.9	215	19.2	152	21.3
$25 000–$34 999	361	19.7	225	20.1	136	19.0
$35 000–$49 999	416	22.7	256	22.8	160	22.4
≥$50 000	434	23.6	268	23.9	166	23.2
Currently married[b]						
Yes	1520	78.4	893	75.8	627	82.5
No	418	21.6	285	24.2	133	17.5
Screening risk group[b]						
Moderate	942	48.4	655	55.2	287	37.7
High	1006	51.6	532	44.8	474	62.3
History of high blood pressure[b]						
Yes	582	30.1	288	24.6	294	38.7
No	1349	69.9	884	75.4	465	61.3
History of heart condition[b]						
Yes	216	11.3	92	8.0	124	16.5
No	1691	88.7	1064	92.0	627	83.5
Smoking status[b]						
Current smoker	282	14.7	189	16.2	93	12.3
Former smoker	734	38.2	407	34.9	327	43.4
Never smoked	904	47.1	570	48.9	334	44.3

[a]Sample size decreased to 1948 as a result of missing data on history of high blood cholesterol; sample sizes were inconsistent because of missing data.
[b]Difference between no history and prior history $P < .000$ by χ^2 test; age, smoking status: $df = 2$; currently married, screening risk group, history of high blood pressure and heart condition: $df = 1$.
[c]Difference between no history and prior history $P < .05$ by χ^2 test, $df = 3$.

ate," or "high" risk; susceptibility to heart attack as a consequence of high cholesterol; and benefits of diet and drug therapy for lowering cholesterol. Participants with elevations were advised to see their doctor for retesting and diagnosis. Handouts containing this information were provided.

2. A letter restating the referral and explaining a coupon offer totaling $17 for heart-healthy foods, redeemable at the sponsoring supermarket for each participant adhering to the referral, was mailed approximately 3 days after screening.

3. The reminder letter reemphasized the referral, described the association between elevated cholesterol and heart attack and the benefits of physician-directed treatments to lower cholesterol, and included a cue to action (screening level and risk group). It was mailed approximately 3 days after screening.

Analysis

Analyses were stratified by history of high blood cholesterol. Bivariate contingency table analyses compared the unadjusted rates of referral compliance between communicator conditions overall and within each second-level intervention group. Multiple logistic regression analyses with backward selection were performed to identify significant covariates for inclusion in the final model; variables included communicator type and behavioral interventions and their interaction. Adjusted odds ratios (ORs) were estimated by exponentiating the maximum likelihood estimates of the parameters (beta coefficients), and asymmetric 95% confidence intervals (CIs) were calculated using the asymptotic standard errors of the coefficient estimates.[11]

After intervention variables and their interactions were entered, the final model included dichotomous variables for marital status, screening cholesterol risk level, and history of high blood pressure and heart condition; cigarette smoking was included as a categorical variable. Continuous variables included were counseling experience, age, obesity (Quetelet's index), and 5-point measures for intention to see a doctor if cholesterol is high (1 = "very likely," 5 = "not likely at all"), perceptions of difficulty to see a doctor (1 = "strongly agree," 5 = "strongly disagree"), and "current health" (summated score of agreement for "somewhat ill," "health is excellent" [reversed scale], and "feeling bad lately"; range = 3 to 15; Cronbach alpha = .73). We used SAS for bivariate analyses and Biomedical Data Processing (BMDP) for logistic regression.

Results

Participant Characteristics

There were no differences in sociodemographic characteristics or coronary heart disease risk factors between counselor types or among the six study groups. There were three sociodemographic differences between history groups: "no history" subjects were younger, somewhat more educated, and less likely to be married (Table 2). As expected, fewer unaware subjects had any particular major coronary heart disease risk factor.

Compared with the adult population of Rochester and its suburban areas, referred subjects were more likely to be White, older, and female and to have higher levels of education.[12] Only 16.2% of unaware and 12.3% of "prior history" subjects were current smokers compared with approximately 25% of both the total screened and Rochester area populations.[13]

Communicator and Behavioral Intervention Comparisons

Within 5 months after screening, 1156 referred subjects (58.5%) visited their physician: 645 (54.8%) of the no history group and 499 (66%) of the prior history group. For no history subjects, communicator type yielded no overall unadjusted effect on the proportion seeking medical care (Table 3). An impact of both behavioral interventions was observed, with the coupon and reminder about equally effective

($P < .0001$). Only the coupon offer enhanced compliance with professional counseling ($P = .02$), while for lay communicators both behavioral interventions were effective ($P < .0001$).

Stratification by behavioral interventions revealed communicator differences only for the reminder: the lay–reminder condition produced the highest level of referral compliance (66.7%, 95% CI = 59.7, 73.6), which was above the upper limit of the 95% CI of the professional–reminder group (50.2%; 95% CI = 43.5, 56.9) ($df = 1$, $P = .001$). Finally, prior history subjects were generally 10% to 15% more likely to have visited a physician, but neither communicator type nor behavioral interventions significantly increased outcome (data not shown).

Adjustment for Covariates

The final adjusted stepwise regression analysis in Table 4 shows that, for unaware subjects, the presence of covariates in the model did not alter the interaction between communicator type and behavioral interventions and differences in compliance rates ($P = .005$). With professional counseling, coupon offer subjects were more likely than control subjects to comply with the referral; the reminder letter was not effective. The behavioral intervention effect was most evident for lay-counseled subjects: both the coupon offer (OR = 2.52, 95% CI = 1.52, 4.18) and the reminder letter (OR = 3.10, 95% CI = 1.83, 5.22) were associated with an increased likelihood of a physician visit.

Older subjects and those with other coronary heart disease risk factors were slightly more likely to have visited a physician. High compared with moderate screening risk level and expression of a greater intention to seek medical care for an elevated test were associated with compliance.

In the model for prior history subjects (data not shown), no differences in seeking medical care were associated with either communicator type (OR = 1.03, 95% CI = 0.73, 1.45) or follow-up interventions (coupon offer: OR = 1.16, 95% CI = 0.77, 1.76; reminder letter: OR = 1.32, 95% CI = 0.87, 2.01). Receipt of counseling from a more experienced communicator, being older, and currently having a spouse or partner emerged as weak predictors, as well as high risk level and a history of high blood pressure. The most powerful predictor was greater intention to seek medical care (OR = 0.70, 95% CI = 0.59, 0.82, $P < .0001$).

TABLE 3—Compliance with High Blood Cholesterol Screening Referral Recommendation: Communicator and Behavioral Intervention Effects for No History Participants

Communicator and Behavioral Intervention Group Comparison	Total No History[a] (n = 1178) n	%	Physician Visit n	%	95% Confidence Interval[b]
Communicator					
Professional	640	54.3	344	53.8	49.9, 57.6
Lay	538	45.7	301	56.0	51.8, 60.1
Behavioral intervention[c]					
Coupon offer	389	33.0	236	60.7	55.8, 65.5
Reminder letter	392	33.3	226	57.7	52.8, 62.5
Control	397	33.7	183	46.1	41.2, 51.0
Communicator and behavioral intervention					
Professional[d]					
Coupon offer	202	31.6	125	61.9	55.2, 68.6
Reminder letter	215	33.6	108	50.2	43.5, 56.9
Control	223	34.8	111	49.8	43.2, 56.3
Lay					
Coupon offer	187	34.8	111	59.4	52.3, 66.4
Reminder letter	177	32.9	118	66.7	59.7, 73.6
Control	174	32.3	72	41.4	34.1, 48.7

[a]Sample size based on respondents with information on physician visit and history of high blood cholesterol.
[b]Confidence intervals are derived from the normal approximation to the binomial distribution.
[c]Difference among behavioral intervention groups $P < .0001$ by χ^2 test, $df = 2$.
[d]Difference among behavioral intervention groups $P = .02$ by χ^2 test, $df = 2$.

To further investigate the interaction in the no history group, alternative unadjusted models, assuming that various combinations of the six study group rates were equal, were evaluated (data not shown). A satisfactory fit was obtained ($\chi^2 = 5.88$, $df = 4$, $P = .21$) for a model that assumed similar compliance rates for the professional and lay coupon groups and the lay–reminder group, as well as the two control groups and the professional–reminder group. The combined rate for the coupon offer and lay–reminder groups was 62.5%, while the other three groups had a rate of 47.6%. A simpler hierarchy of follow-up interventions resulted compared with the observed data (Table 3).

There were no differences in rates of participants seeking care due to variations in effectiveness among counselors. Individual counselor rates for no history subjects were inconsistent across follow-up intervention groups. Removal of three lay counselors with a greater proportion of compliant referrals in the reminder condition diminished the lay–reminder effect; however, both behavioral interventions maintained higher visit rates (lay–coupon offer = 62.5% and lay–reminder letter = 59.3% vs lay–control = 40.5%; $P < .0001$). Finally, differences in participant characteristics in the lay–reminder group did not contribute to the interaction.

Discussion

A 1990 national survey estimates that 29 million US adults have attended public cholesterol screenings.[14] The inadequacy of counseling and lack of physician referral has been cited by the Office of the Inspector General.[15] The present study, which contributed to the NHLBI Workshop on Public Screening for Blood Cholesterol recommendations,[16] provides support for practical counseling and follow-up strategies that can be adopted by other screening programs to increase the chances of those referred seeking appropriate retesting and diagnosis.

Screening with immediate cholesterol assessment provides a teachable moment in which participants are highly receptive to coronary heart disease and dietary education and referral messages.[17] Findings substantiate the hypothesis of no differences in counseling effectiveness between professional and lay communicators. Consistent with other studies showing that lay individuals performed comparably to health professionals in influencing health attitudes and behaviors,[18,19] these results have implications for those planning screening. Lay counselors are more accessible and less costly to train and employ than health profes-

TABLE 4—Logistic Regression Model of Factors Associated with High Blood Cholesterol Screening Referral Compliance for No History Participants[a]

Variable	Adjusted Odds Ratio[b]	95% Confidence Interval	*P* Value
Communicator and behavioral intervention interaction			.005
Professional			
Coupon[c]	1.94	1.21, 3.09	.005
Reminder letter[c]	1.04	0.67, 1.63	.840
Lay			
Coupon[c]	2.52	1.52, 4.18	.001
Reminder letter[c]	3.10	1.83, 5.22	.0001
Sociodemographics			
Age[d]	1.19	1.05, 1.34	.005
Coronary heart disease risk factors			
Screening cholesterol risk level[e]	2.97	2.23, 3.96	.0001
History of heart condition[f]	1.64	0.96, 2.79	.066
History of high blood pressure[f]	1.34	0.95, 1.90	.095
Obesity	0.97	0.95, 0.99	.015
Attitudes/beliefs			
Intention to visit physician	0.53	0.45, 0.62	.001
Barrier to physician visit	1.21	1.07, 1.37	.009
Current Health Index	0.92	0.87, 0.97	.003

[a]Sample size decreased to 1002 as a result of missing data on covariates (68 related to family income).
[b]Odds ratios and confidence intervals for fixed communicator and behavioral intervention factors estimated from the logistic regression analysis.
[c]Odds ratio compares intervention vs control.
[d]Odds ratio compares risk of 10-year difference (e.g., 20–29 years).
[e]Odds ratio compares high vs moderate.
[f]Odds ratio compares yes vs no.

sionals, and extra resources can be directed to referral follow-up.

The results suggest that the influence of the behavioral interventions depends on cholesterol history and communicator type. For those with newly identified elevations, both types of follow-up yielded better physician visit rates than no follow-up. Although the rates were higher than found in studies without follow-up only for previously diagnosed subjects, direct comparisons are limited since these reports are not adjusted for prior awareness and rely on small samples with low response rates.[5,6]

The higher compliance rates for the coupon intervention contribute to the promising use of material incentives in other coronary heart disease risk programs.[20–22] However, questions of "how much" incentive is necessary and whether $17 in redeemable coupons for low-cholesterol foods would produce a similar effect in minority and lower socioeconomic populations need to be investigated. The reminder letter emphasized health beliefs associated with undertaking recommended health behaviors,[23] a more effective approach than reminder messages from a physician.[24] This may account for the lack of significant referral outcomes with the use of a physician follow-up letter by Lefebvre and colleagues.[25] The finding that communicator type may contribute to the success of reminders extends results on their effectiveness in cancer detection and influenza-vaccination programs[24,26] and hypertension screening follow-up.[27,28]

Since public screenings attract participants with a broad range of characteristics, it is encouraging that for no history subjects the follow-up interventions were more important than participant characteristics and attitudes in the adjusted analysis. The exceptions were risk level and behavioral intention (predictive of health behavior in earlier studies),[29] suggesting inclusion of these issues in counseling and follow-up efforts. The equivocal effect of the reminder letter by communicator type is not supported by prior studies or motivational theory. Referral was based on age-specific cutpoints widely in use at the time of the screenings.[2] The present guidelines recommend referral at lower cholesterol levels (total serum cholesterol: 200 to 239 mg/dL) when two or more coronary heart disease risk factors other than high cholesterol are present.[3] It is not known whether our findings will apply equally at these lower levels, but the regression suggests that referrals at the new lower cutoffs may be less likely to seek retesting and diagnosis and would benefit even more from follow-up.

The logistic regression models should be interpreted with caution, as all model-selection methods are subject to error, and the reasonableness of assumptions can never be completely checked.[30] In a simple assessment of the predictive value of the final model by calculating the probability of a physician visit for each subject, comparison of the observed rates of compliance by no history subjects with the fitted estimates showed misclassification in only 12.8% of those seeking care ($P < .50$) and 17.7% of noncompliers ($P > .50$).

Although attrition between screening and follow-up was only 7% (and equally distributed among study groups), reanalyses of unadjusted results, conservatively considering "dropouts" to be "noncompliers," did not change the outcomes for no history subjects.

Barriers to screening follow-up due to enabling factors[31] did not occur (data not shown). Almost all referrals reported having a physician and health insurance or entitlement coverage, and no differences occurred in seeking medical follow-up by varying reimbursement for preventive services (e.g., private health plan, health maintenance organization, Medicaid). For unaware subjects, the most frequent reason for not seeking care was attributed to ability to self-manage high cholesterol (58.3%), and 45.8% of prior history subjects reported already being in treatment. Less than 6% of either group mentioned the lack of a physician, and treatment costs were a deterrent for less than 1%.

Investigation of referral compliance in the context of public screenings poses some limitations. Participants self-selected to attend and may have been more health conscious than the general Rochester population. The follow-up interventions might not produce similar incremental effects in programs where attendance stems less from individual initiative. Overestimation in rates due to physician visits for reasons unrelated to the referral may have occurred (e.g., some previously diagnosed individuals attended proximal to a scheduled "checkup" and 13% of nonreferred "usual" risk participants reported a post-screening visit), but the relative comparisons are valid in the context of the experimental design. Finally, although these findings cannot be generalized to minority populations, there are similarities to the Boston Model Systems Project, which had a higher proportion of minority attendance.[32]

The concern that public screenings will reach those already aware of their condition becomes increasingly important as more people report having had a cholesterol test (from 35% in 1983 to 65% in 1990).[14] Only 31% of aware participants were currently on a cholesterol-lowering regimen, suggesting that it may be beneficial to target those who have discontinued treatment to reinitiate appropriate medical care.

From a programmatic perspective, both types of follow-up interventions appear reasonably effective for newly identified participants. The reminder letter, with a cost of approximately $1.50 per subject for staff time and mailing, produced a similar percentage increase (18% adjusted) to the coupon offer but was only effective with lay communicators. Although the coupon offer resulted in generally better visit rates, the efforts and costs for redemption and distribution in relation to the 12% to 18% incremental effects should be considered.

Millions of individuals have been screened in public settings throughout the United States and will continue to attend these programs. The present study suggests some practical strategies to assist in linking referrals to medical care. □

Acknowledgments

The research was supported by Grant HL37895 from the National Heart, Lung, and Blood Institute. The first author was supported by Research Career Development Award HD00538 from the National Institute of Child Health and Human Development, and the last author by Preventive Cardiology Academic Award HL00982 from the National Heart, Lung, and Blood Institute.

This paper was presented in part at the 2nd Annual Primary Care Conference, Agency for Health Care Policy and Research, San Diego, Calif, January 1991.

We gratefully acknowledge the assistance of Kathryn L. Doane, MPH, for coordinating the project and data collection; Nancy L. Bowley, RN, for supervising the screenings; and Joseph J. Guido, MS, for programming. Appreciation is expressed to Mary Ellen Burris and Wegmans Food Markets Inc for assistance in implementing the project.

References

1. Cleeman J, Lenfant C. New guidelines for the treatment of high blood cholesterol in adults from the National Cholesterol Education Program. *Circulation.* 1987;76:960–962.
2. National Institutes of Health Consensus Development Conference. Lowering blood cholesterol to prevent heart disease. *JAMA.* 1985;253:2080–2086.
3. National Cholesterol Education Program. Report of the expert panel on population strategies for blood cholesterol reduction: executive summary. *Arch Intern Med.* 1991;151:1071–1084.
4. Wynder EL, Field F, Haley NJ. Population screening for cholesterol determination: a pilot study. *JAMA.* 1986;258:2839–2842.
5. Rastam L, Luepker RV, Pirie PL. Effect of screening and referral on follow-up and treatment of high blood cholesterol levels. *Am J Prev Med.* 1988;4:244–248.
6. Bradford RH, Bachorik PS, Roberts K, Williams OD, Gotto AM. Blood cholesterol screening in several environments using a portable, dry chemistry analyzer and fingerstick blood samples. *Am J Cardiol.* 1990;65:6–13.
7. Greenland P, Levenkron JC, Radley MG, Baggs JG, Manchester RA, Bowley NL. Feasibility of large-scale cholesterol screening: experience with a portable capillary-blood testing device. *Am J Public Health.* 1987;77:73–75.
8. Maiman LA, Greenland P, Hildreth NG, Cox C. Patterns of physicians' treatments for referrals from public cholesterol screening. *Am J Prev Med.* 1991;7:273–279.
9. Greenland P, Bowley NL, French CA, Meiklejohn B, Gaghano S, Sparks CE. Precision and accuracy of portable blood analyzer system during cholesterol screening. *Am J Public Health.* 1990;80:181–184.
10. National Cholesterol Education Program. Current status of blood cholesterol measurement in clinical laboratories in the United States: a report from the Laboratory Standardization Panel. *Clin Chem.* 1988;34:193–201.
11. Schlesselman JJ. *Case Control Studies: Design, Conduct, Analysis.* New York, NY: Oxford University Press;1982:233–248.
12. US Bureau of the Census. *1980 Census of Population and Housing: Census Tracts Rochester, New York.* Washington, DC: US Government Printing Office;1983.
13. Giovino GA, Ossip-Klein D, Marin G, Valenzuela LM. An assessment of 1987 Monroe County Clean Indoor Air legislation restricting public smoking. Presented at the Upstate New York and Southern Ontario Epidemiology Meeting; October 1987; Albany, NY.
14. Schucker B, Wittes JT, Santanello NC, et al. Change in cholesterol awareness and action: results from national physician and public surveys. *Arch Intern Med.* 1991;151:666–673.
15. Office of the Inspector General. *Public Cholesterol Screening.* Washington, DC: Office of Analysis and Inspections; 1989. Publication OAI-05-89-01330.
16. National Heart, Lung, and Blood Institute. Recommendations regarding public screening for measuring blood cholesterol: summary of a National Heart, Lung, and Blood Institute workshop, October 1988. *Arch Intern Med.* 1989;149:2650–2654.
17. Van Beurden E, James R, Dunn T, Tyler C. Risk assessment and dietary counseling for cholesterol reduction. *Health Educ Res Theory Pract.* 1990;5:445–450.
18. Wingert WA, Larson W, Friedman DB. Indigenous health aides as counselors to parents about nutrition. *Public Health Rep.* 1969;84:328–332.
19. Holder L. Some theoretical and practical considerations in influencing health behavior. *Health Educ Monogr.* 1970;30:49–67.
20. Brengelmann J, Sedlmayr E. Experiments in the reduction of smoking behavior. In Steinfied et al., eds. *Health Consequences, Education, Cessation Activities and Governmental Action.* Washington, DC: US Government Printing Office; 1977. DHEW publication no. (NIH) 77-1413.
21. Brownell K, Cohen R, Stunkard A, Felix M, Cosby N. Weight loss competition at the worksite: impact on weight, morale and cost-effectiveness. *Am J Public Health.* 1984;74:1283–1285.
22. Klesges R, Vasey M, Glasgow R. A worksite smoking modification competition: potential for public health impact. *Am J Public Health.* 1986;76:198–200.
23. Janz NH, Becker MH. The Health Belief Model a decade later. *Health Educ Q.* 1984;11:1–47.
24. Larson EB, Bergman J, Heindrich F, Alvin BL, Schneweiss R. Do postcard reminders improve influenza vaccination compliance? A prospective trial of different postcard "cues." *Med Care.* 1982;20:639–648.
25. Lefebvre RC, Banspach SW, Gans KM, Carleton RA, Lasater TM. Enhancing adherence to referral advice given at blood cholesterol screenings: impact on participant follow-up and physician behavior. *Health Educ Res Theory Pract.* 1991;6:405–413.
26. Fink R, Shapiro S, Roesser R. Impact of efforts to increase participation in repetitive screening for early breast cancer detection. *Am J Public Health.* 1972;62:328–336.
27. Velez R, Anderson L, McFall S, Magruder-Habib K. Improving patient follow-up in incidental screening through referral letters. *Arch Intern Med.* 1985;145:2184–2187.
28. Murray DM, Kurth CL, Finnegan JR, Pirie PL, Admire JB, Luepker RV. Direct mail as a prompt for follow-up care among persons at risk for hypertension. *Am J Prev Med.* 1988;4:331–335.
29. Cummings KM, Jette AM, Brock BM, Haefner DP. Psychosocial determinants of immunization behavior in a swine influenza campaign. *Med Care.* 1979;17:639–649.
30. Greenland S. Modeling and variable selection in epidemiologic analysis. *Am J Public Health.* 1989;79:340–349.
31. Zapka JG, Stoddard AM, Costanza ME, Greene HL. Breast cancer screening by mammography: utilization and associated factors. *Am J Public Health.* 1989;79:1499–1502.
32. Havas S, Koumjian L, Reisman J, Hsu L, Wozenski S. Results of the Massachusetts model systems for blood cholesterol screening project. *JAMA.* 1991;266:375–381.

Food Sources, Dietary Behavior, and the Saturated Fat Intake of Latino Children

Charles E. Basch, PhD, Steven Shea, MD, and Patricia Zybert, PhD

ABSTRACT

Background. Recent recommendations for Americans aged 2 and older call for a reduction in the average saturated fat intake to less than 10% of calories.

Methods. Using 24-hour dietary recalls collected from mothers of 4- to 7-year-old urban Latino children, we identified foods and dietary behavior patterns that distinguish children with higher and lower mean daily percentages of calories from saturated fat.

Results. Compared with children in the lowest quintile of intake, children in the highest quintile consumed more than twice as much saturated fat per day from high-fat milk products (18.5 g vs 7.8 g), mostly from whole milk. They did not consume different kinds of milk or different amounts of milk per eating occasion, but on average they consumed milk more frequently (2.8 vs 1.6 eating occasions per day). Even children in the lowest quintile, on average, exceeded the 10% of calories from saturated fat currently recommended. If low-fat (1% fat) milk had been substituted without other dietary changes, all but the highest two quintiles would have been within the recommended level.

Conclusions. The substitution of low-fat for whole milk appears to be a key strategy for preschool children for achieving recommended levels of saturated fat intake. (*Am J Public Health.* 1992;82:810–815)

Introduction

Numerous studies in adults link diet to risk of cardiovascular disease, some types of cancer, and other illness.[1–3] Dietary saturated fat in particular has been found to be associated with atherosclerosis.[2] Studies in young children[4–7] have found that total and saturated fat intakes are higher than recommended levels,[8–10] that saturated fat is associated with serum total and low-density lipoprotein cholesterol levels,[11] and that there is tracking of serum lipid[12] and nutrient intake levels.[6,13,14] One risk reduction objective for the year 2000, specified by the US Public Health Service (PHS)[10] and supported by the National Cholesterol Education Program (NCEP),[9] is to reduce average saturated fat intake among children aged 2 years and older from more than 13% to less than 10% of total calories.

We identified only two studies that examined the contributions of specific foods and food groups to young children's saturated fat intake.[4,15] The Nationwide Food Consumption Survey is a continuing national survey of food intakes by individuals that uses probability sampling and repeated dietary recall interviews of parents to provide surrogate reports of children's dietary intake. (For 1985, this survey included 222 children aged 1 to 3 and 149 children aged 4 to 5)[4] In this survey, milk products contributed almost 40% of children's saturated fat intake. In the Bogalusa study,[15] which used a single surrogate dietary recall interview per year, milk alone contributed more than 20% of saturated fat intake for children aged 3 (n = 106) and 4 (n = 219).

Conceptualization and measurement of diet can be addressed from multiple perspectives. Most nutritional epidemiology has examined the relationship between nutrient intake levels and disease or risk factors.[1–3] Several recent reports demonstrate the value of analyzing diet in terms of dietary behavior patterns, such as the amount and type of food eaten[16] and the frequency of eating throughout the day.[17] According to Willett,[18] educational interventions to assist the public in making healthful food choices should focus on specific foods and dietary patterns rather than on nutrients. To our knowledge, no published research has described the dietary patterns that determine children's saturated fat intake.

We sought to determine whether children with higher and lower saturated fat intakes eat different foods, or eat the same foods but have different dietary behavior patterns (i.e., do they eat different amounts per serving or eat the same serving sizes but eat more frequently). To do this, we collected 24-hour dietary recalls for young Latino children on seven occasions over 3 years. These data were used to determine the relative contributions of food sources to children's saturated fat intake, and to identify foods and dietary behavior patterns that distinguish children with higher and lower saturated fat intakes.

Charles E. Basch and Patricia Zybert are with the Center for Health Promotion, Teachers College, Columbia University, and Steven Shea is with the College of Physicians & Surgeons and School of Public Health, Columbia University, in New York City.

Requests for reprints should be sent to Charles E. Basch, Center for Health Promotion, Box 114, Teachers College, Columbia University, 525 West 120th Street, New York, NY 10027.

This paper was submitted to the Journal June 25, 1991, and accepted with revisions November 27, 1991.

Methods

Study Sample

Subjects were 4- to 7-year-old Latino children and their mothers living in northern Manhattan or the Bronx, New York City, and participating in the Columbia University Study of Childhood Activity and Nutrition. Ten percent of the mothers and more than 90% of the children for whom we had birthplace data (n = 204 mothers and n = 184 children) were born in the United States. Children were generally healthy, with no dietary prescriptions or restrictions. Recruitment, which was mainly from a pediatric practice of a large hospital in a disadvantaged urban community, is described in greater detail elsewhere.[11] All mother/child pairs who completed four 24-hour recalls from the first year of the study were included in the baseline analysis (n = 205 pairs; 103 male children). Of these pairs, 168 completed three 24-hour dietary recalls in the third year of the study as well. At the time of first recall, the children's mean age was 4.4 years (SD = 0.34) and the mother's mean age was 31.8 years (SD = 5.2). At the time of the fifth recall, the children's mean age was 6.3 years (SD = 0.42).

Data were elicited orally by trained interviewers from 1986 to 1989. Because Spanish was the language spoken at home by 95% (n = 195) of these families most of interviews were conducted in Spanish.

Measurements

During four fairly evenly spaced visits to our community field site in the first year and three visits in the third year of the study, each mother recalled her child's dietary intake during the previous 24 hours. After listing the foods eaten by her child during each eating occasion, she was then asked to describe portion sizes. A variety of probing techniques and adjuncts, such as three-dimensional food models, were used to enhance the completeness and accuracy of the mothers' reports. Dietary data were cleaned, coded, and edited by nutritionists prior to data entry and verification. Nutrient values were computed using the Georgetown University Study of Childhood Activity and Nutrition dietary analysis program.[19] Nutrient values for commonly eaten foods that were not included in the original database were added.[20–22] Previously reported data provide support for the reliability[6] and validity[11,23,24] of the mothers' reports of their children's diet.

Statistical Analysis

Typical nutrient intakes at two points in time were derived by averaging results from the four 24-hour recalls collected in year 1 (n = 205) and from the three recalls collected in year 3 (n = 168). Mean number of eating occasions, quantity, and grams of saturated fat from each food were calculated for each child. Foods were collapsed into 33 food groups (Appendix 1) based on a typology adapted from the US Department of Agriculture.[21]

Using data from year 1, a stepwise multiple regression analysis was used to select those foods that contributed the most to variance in percentage of calories from saturated fat. The possible independent variables were grams of saturated fat consumed per day from each of the 33 food groups. Grams of saturated fat for selected specific foods, most notably whole milk, were also examined.

Children were classified according to percentage of caloric intake from saturated fat and divided into quintiles. Joint classification analysis determined the stability of the quintile classification from year 1 to year 3.[25] More than 60% of the children who were classified in the highest (or lowest) quintile in year 1 remained in the highest (or lowest) two quintiles in year 3. Seasonal variation was not associated with saturated fat intake.

Stepwise discriminant analyses were used to select (1) those food groups whose grams of saturated fat per day and (2) those food groups whose number of servings per day most accurately distinguished children in the lowest and highest quintiles. The analysis by servings allowed us to identify foods both with and without fat that were associated with saturated fat intake. Prespecified serving sizes for each food were taken from the nutrient database (e.g., one cup of whole milk or one ounce of cheese = one serving). Multivariate analysis of variance was used to assess the overall significance of differences in the selected food groups across all five quintiles. Mean number of servings and mean number of grams of saturated fat per day were tested for univariate linear trends across the quintiles. The selected food groups were examined more closely to identify dietary behavior patterns, including specific foods mentioned, average amount consumed per eating occasion, number of eating occasions per day, and type of eating occasion (i.e., whether foods were consumed as part of snacks or main meals).

Data from year 3 were then used to verify the results. Mean percentage of calories from saturated fat was predicted using the food groups selected by the stepwise regression on year 1 data. The discriminant function generated from year 1 data was calculated using year 3 intakes and was used to classify subjects into extreme quintiles of percentage of calories from saturated fat in year 3.

All the analyses were replicated using calorie-adjusted saturated fat instead of percentage of calories from saturated fat. By using residuals from a regression with caloric intake as the independent variable and with saturated fat intake as the dependent variable, inferences can be drawn about saturated fat intake independent of total caloric intake.[3,25,26] Results were the same for calorie-adjusted saturated fat and nutrient densities. Only data for percentage of calories from saturated fat are reported because current recommendations are expressed in this way.[8–10]

Results

Data in Table 1 show that the reported mean percentage of calories from saturated fat in year 1 was 13.3% (SD = 2.2), approximately 30% higher than the amount of calories (less than 10%) recommended by the PHS.[9,10] Even children in the lowest quintile had an average reported consumption of more than 10% of total calories from saturated fat. High-fat milk products accounted for more than half and whole milk alone contributed 44% of the total reported saturated fat consumption across the sample.

Reported grams of saturated fat consumed per day from the following five food groups made the greatest successive incremental contributions to explained variance in the children's percentage of calories from saturated fat: high-fat milk products (R^2 = .56), beef/veal (R^2 = .60), enriched bread (R^2 = .62), animal fats/oils for cooking and table (R^2 = .65), and pork/lamb/sausage (R^2 = .67). Whole milk alone accounted for 41% of the variance.

Compared with children in the lowest quintile, children in the highest quintile consumed more than twice as much saturated fat per day from high-fat milk products (18.5g versus 7.8g), including more than twice as much from whole milk (13.2g vs 6.4g) and more than six times as much from hard cheese (3.4 vs 0.5g). The amounts of saturated fat from cheese were much smaller than those from whole milk.

TABLE 1—Daily Saturated Fat Intake from Selected Food Groups and All Foods, Calories from Saturated Fat, and Total Energy Intake by Quintile of Children's Percentage of Calories from Saturated Fat

	Lowest Quintile (n = 41)		Second Lowest Quintile (n = 41)		Middle Quintile (n = 41)		Second Highest Quintile (n = 41)		Highest Quintile (n = 41)		Total Sample (n = 205)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
High-fat milk products (g)	7.8	(3.2)	10.3	(2.7)	12.2	(3.0)	14.7	(2.7)	18.5	(5.1)***	12.7	(5.1)
Whole milk	6.4	(3.0)	8.4	(3.0)	10.0	(2.7)	12.2	(2.6)	13.2	(4.3)***	10.0	(4.0)
Hard cheese	0.5	(0.9)	1.1	(1.5)	1.7	(1.7)	1.7	(1.7)	3.4	(3.0)***	1.7	(2.1)
Other high-fat milk products	0.9	(1.0)	0.8	(1.2)	0.5	(0.7)	0.7	(1.0)	2.0	(4.3)*	1.0	(2.2)
Beef/veal (g)	0.8	(1.0)	1.5	(2.3)	1.4	(1.5)	1.8	(1.8)	1.9	(2.1)**	1.5	(1.8)
Enriched bread, etc. (g)	0.4	(0.2)	0.4	(0.2)	0.4	(0.2)	0.4	(0.3)	0.4	(0.2)	0.4	(0.2)
Animal fats/oils for cooking and table (g)	0.4	(0.8)	0.4	(0.7)	0.4	(0.6)	0.5	(0.9)	0.6	(1.4)	0.4	(0.9)
Pork/lamb/sausage (g)	0.1	(0.5)	0.4	(0.9)	0.6	(1.5)	0.5	(0.9)	0.4	(1.0)	0.4	(1.0)
Saturated fat (all foods) (g)	17.0	(4.0)	21.4	(6.1)	22.0	(5.0)	24.7	(3.7)	29.9	(8.4)***	23.0	(7.1)
Calories saturated fat (%)	10.3	(1.0)	12.1	(0.4)	13.2	(0.3)	14.5	(0.3)	16.5	(1.1)***	13.3	(2.2)
Total energy intake (calories)	1481.8	(290.8)	1588.4	(436.1)	1495.6	(327.3)	1539.7	(227.3)	1636.3	(466.9)	1548.4	(362.5)

*$P < .05$ linear trend.
**$P < .01$ linear trend.
***$P < .001$ linear trend.

Parents of only four children in year 1 reported the consumption of low-fat or skim milk.

A discriminant function using grams of saturated fat per day from high-fat milk products and from beef/veal correctly classified 95% of the children in the two most extreme quintiles. A discriminant function using only grams of fat from whole milk correctly classified 84%. The multivariate analysis of variance showed significant overall differences among the five quintiles ($\Lambda = .43$, $P < .001$). Univariate tests for linear trends in grams of saturated fat from high-fat milk products and from beef/veal were both significant ($P < .001$ and $P < .01$, respectively).

A discriminant function using reported mean servings per day of high-fat milk products and of soda/punch/fruit drinks correctly classified 90% of the children in the two most extreme quintiles of percentage of calories from saturated fat. More reported servings per day of high-fat milk products and fewer servings per day of soda/punch/fruit drinks were associated with a high percentage of calories from saturated fat (Table 2). A discriminant function using only servings of whole milk correctly classified 84% of the children. The multivariate analysis of variance on number of servings per day indicated significant overall differences among the five quintiles ($\Lambda = .56$, $P < .001$). Univariate tests for linear trends in servings of high-fat milk products and of soda/punch/fruit drinks were both significant ($P < .001$) (Table 2).

Specific dietary behavior patterns involving whole milk were examined. In year 1, the mean number of servings per day ranged from 1.3 cups in the lowest quintile to 2.6 cups in the highest quintile (Table 2). The children in the lowest quintile did not consume smaller amounts of whole milk per eating occasion than those in the highest quintile (0.8 cups vs 0.9 cups) but consumed whole milk less frequently (1.6 vs 2.8 times per day). Overall, 56% of whole milk was consumed as part of snacks: 48% for children in the lowest quintile and 62% for children in the highest quintile. Thus, a substantial portion of whole milk was consumed daily as snacks, regardless of quintile membership.

Data from year 3 confirm the patterns observed in year 1. Grams of saturated fat consumed in year 3 from the five food groups selected using data from year 1 accounted for 58% of the variance in percentage of calories from saturated fat for year 3 data. The discriminant function from year 1 calculated using grams of saturated fat from year 3 correctly classified 79% of the children in the two most extreme quintiles in year 3. The multivariate analysis of variance on grams of saturated fat across quintiles was significant ($\Lambda = .58$, $P < .001$), but only the univariate trend for grams of saturated fat from high-fat milk products remained significant ($P < .001$). The discriminant function generated from year 1 data calculated with daily servings of high-fat milk products and soda/punch/fruit drinks in year 3 correctly classified 74% of the children in the two most extreme quintiles of percentage of calories from saturated fat in year 3 (Table 2). The multivariate analysis of variance on servings across quintiles was significant ($\Lambda = .74$, $P < .001$), but only the univariate linear trend for servings of high-fat milk products remained significant ($P < .001$). Although reported consumption of whole milk was lower in year 3 for all quintiles, children in the lowest quintile still consumed approximately half the number of cups per day as children in the highest quintile.

Discussion

The PHS[10] and the NCEP[9] recommend that children aged 2 years and older should consume less than 10% of their total calories as saturated fat. In our study, mean saturated fat consumption was 13.3% of total calories, and even children in our lowest quintile of saturated fat intake had reported an average consumption of more than 10%. Consistent with other studies,[4,15] we found that milk and milk products were the foods that made the greatest contribution to young children's saturated fat intake. Our data show that frequent consumption of whole milk was the single most important dietary be-

TABLE 2—Mean Number of Servings per Day of Selected Food Groups from Four Year 1 Recalls and from Three Year 3 Recalls by Quintile of Children's Percentage of Calories from Saturated Fat

	Year 1[a]									
	Lowest Quintile (n = 41)		Second Lowest Quintile (n = 41)		Middle Quintile (n = 41)		Second Highest Quintile (n = 41)		Highest Quintile (n = 41)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
High-fat dairy	1.6	(0.9)	2.0	(0.6)	2.4	(0.7)	2.8	(0.5)	3.4	(1.0)*
Whole milk	1.3	(0.6)	1.6	(0.6)	2.0	(0.5)	2.4	(0.5)	2.6	(0.8)*
Soda/punch/fruit drinks	0.8	(0.8)	0.7	(0.6)	0.7	(0.6)	0.5	(0.3)	0.5	(0.4)*
	Year 3[b]									
	Lowest Quintile (n = 33)		Second Lowest Quintile (n = 34)		Middle Quintile (n = 34)		Second Highest Quintile (n = 34)		Highest Quintile (n = 33)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
High-fat dairy	1.2	(1.1)	1.5	(0.9)	1.9	(1.0)	2.2	(0.7)	2.7	(1.0)*
Whole milk	0.9	(0.6)	1.2	(0.5)	1.6	(0.7)	1.6	(0.7)	2.0	(0.7)*
Soda/punch/fruit drinks	0.6	(0.5)	0.7	(0.7)	0.5	(0.5)	0.7	(0.7)	0.5	(0.5)

[a]Overall $\Lambda = .56$; $df = 8,398$; $P < .001$.
[b]Overall $\Lambda = .74$; $df = 8,324$; $P < .001$.
*Univariate linear trend $P < .001$.

havior pattern that distinguished children with low and high saturated fat intake.

Substitution of 1% milk-fat milk for whole milk, without changing any other dietary behaviors, would have reduced the percentage of calories from saturated fat by 25% for the whole sample (from 13.3% to just under 10.0%) and by no less than 20% in any of the five quintiles. This single dietary substitution would have brought average intakes of all children but those in the two highest quintiles within the level recommended by the PHS[10] and the NCEP.[9]

Concerns have been raised about adverse effects of interventions to reduce children's fat intake,[27–29] but there are few data that empirically address how a moderately reduced fat diet affects growth or development. One study, based on a small sample of children (n = 40), reported growth failure in eight children associated with reduced caloric and zinc intake resulting from unsupervised dietary treatment of hypercholesterolemia.[29] Most research on the role of caloric intake in children's growth and development has been conducted in developing countries.[30–32] We did not identify any published population-based studies documenting adverse effects on growth or development resulting from a moderately reduced fat diet such as that recommended by the NCEP[9] or the PHS.[10] In our data, children in the lowest quintile of percentage of calories from saturated fat did not differ on average from children in any other quintile on measures of height, weight, body mass index, or sum of skinfolds.[33] (Methods for these measures are reported elsewhere).[34] A related issue is that selectively limiting foods that are high in fat, such as dairy products, may also restrict intake of protein and micronutrients.[27] However, given that low-fat, skim, and whole milk are equivalent sources of protein, carbohydrate, calcium, and most other micronutrients (Table 3), substituting low-fat for whole milk affects only intake of calories and fat. Moreover, balancing these concerns about possible unintended adverse effects of reducing children's calorie and fat intakes are national data showing that prevalence rates of obesity among children are increasing[35] and that many children eat too much saturated fat.[4–7]

Several limitations of this study must be considered. First, data in Table 2 show that the average number of servings of milk and milk products was lower in year 3 than in year 1. We have reported elsewhere that this decline is most likely an artifact of children entering school and mothers thus being unaware of their children's food consumption during lunch.[6] Because this change occurred in all five quintiles, it did not bias the results. Second, we defined saturated fat generically, but some reports indicate that different kinds of saturated fat affect serum cholesterol differently.[36,37] Although further research may lead to more specific recommendations, current dietary recommendations for children by the PHS[10] and the NCEP[9] are specified in these terms.

The generalizability of our findings may be limited because this study included a single sample of 4- to 7-year-old Latino children from low socioeconomic families in one disadvantaged urban community. The Latino population in disadvantaged urban settings is of particular interest for public health because it is growing rapidly[38] and because, compared with non-Hispanic Whites, Latinos have been shown to have increased prevalence of some cardiovascular disease risk factors,[39–44] to have lower levels of knowledge about cardiovascular disease risk factors,[39,44] and to consume more saturated fat and less low-fat and non-fat milk than other ethnic groups.[45,46] Generalizability may also be limited because our data were collected from 1986 to 1989. Over the last two decades, national per capita consumption of whole milk has declined while consumption of low-fat milk has increased,[47] yet less than 4% of the

TABLE 3—Nutrition Information for One Cup Serving of Grade A Pasteurized, Homogenized Whole Milk, 2% Milk-fat Milk, 1% Milk-fat Milk, and Skim Milk

	Whole Milk Vitamin D Added	2% Milk-fat Milk Vitamin A and D Added	1% Milk-fat Milk Vitamin A and D Added	Skim Milk Vitamin A and D Added
Calories	150	120	100	85
Protein (g)	8	8	8	8
Protein[a]	20	20	20	20
Carbohydrate (g)	11	12	12	12
Fat (g)	8	5	3	0
Sodium (mg)	120	122	123	126
Potassium (mg)	370	377	381	406
Vitamin A[a]	4	10	10	10
Vitamin C[a]	4	4	4	4
Vitamin D[a]	25	25	25	25
Vitamin B4[a]	0	0	0	4
Vitamin B6[a]	4	4	4	0
Vitamin B12[a]	15	15	15	15
Calcium[a]	30	30	30	30
Iron[b]				
Thiamine[a]	6	6	6	6
Riboflavin[a]	25	25	25	25
Niacin[b]				
Phosphorus[a]	20	20	20	20
Magnesium[a]	8	8	8	8
Zinc[a]	4	4	4	4
Pantothenic acid[a]	6	6	6	6

[a]Percentage of US daily allowances of these nutrients.
[b]Less than 2% of US recommended daily allowance for these nutrients.

parents of the children in our sample reported consumption of low-fat or skim milk. While our data suggest no substantial shift to low-fat milk, we cannot be sure that this change has not occurred since 1989.

The gap between national estimates showing that 4- to 5-year-old children consume an average of 14.7% of calories from saturated fat[4] and recommendations by the PHS[10] and the NCEP[9] that children aged 2 years and older should consume less than 10% of calories from saturated fat indicates a need for many young children to modify their typical diet. We conclude that, for those young children who consume more than 10% of calories from saturated fat and whose family and pediatrician have decided to pursue dietary change to reach national dietary recommendations,[9,10] an important strategy is for parents to substitute low-fat milk for high-fat milk.[48] Maintenance of sufficient caloric intake should be part of any dietary recommendation. □

Acknowledgments

Supported in part by grants from the National Institutes of Health (HL35189) and the Cancer Research Foundation of America.

References

1. US Department of Health and Human Services. *The Surgeon General's Report on Nutrition and Health.* Washington, DC: US Govt Printing Office; 1988. DHHS PHS publication 88–50210.
2. Committee on Diet and Health, National Research Council. *Diet and Health: Implications for Reducing Chronic Disease Risk.* Washington, DC: National Academy Press; 1989.
3. Willett WC. *Nutritional Epidemiology.* New York, NY: Oxford University Press; 1990.
4. US Department of Agriculture. *Nationwide Food Consumption Survey.* Hyattsville, Md: US Dept of Agriculture; 1987. Human Nutrition Information Service NFCS, CSFII, Report No. 85–4.
5. Kimm SYS, Gergen PJ, Malloy M, Dresser C, Carroll M. Dietary patterns of US children: implications for disease prevention. *Prev Med.* 1990;19:432–434.
6. Stein AD, Shea S, Basch CE, Contento IR, Zybert P. Variability and tracking of nutrient intakes of preschool children based on multiple administrations of the 24-hour dietary recall. *Am J Epidemiol.* 1991;134: 1427–1437.
7. Nicklas TA, Farris RP, Major C, et al. Dietary intakes. *Pediatrics* 1987;80(suppl): 797–806.
8. Task Force of the Nutrition Committee and the Cardiovascular Disease in the Young Council of the American Heart Association. Diet in the healthy child. *Circulation* 1983;67:411A–4A.
9. US Public Health Service. *National Cholesterol Education Program: Report of the Expert Panel on Population Strategies for Blood Cholesterol Reduction Executive Summary.* Washington, DC: US Govt Printing Office; 1990. National Institutes of Health publication 90–3047.
10. US Public Health Service. *Healthy People 2000: National Health Promotion and Disease Prevention Objectives.* Washington, DC: US Govt Printing Office; 1991. US Dept of Health and Human Services publication PHS 91–50212.
11. Shea S, Basch CE, Irigoyen M, et al. Relationships of dietary fat consumption to serum total and low-density lipoprotein cholesterol in Hispanic preschool children. *Prev Med.* 1991;20:237–249.
12. Webber LS, Srinivasan SR, Wattigney WA, Berenson GS. Tracking of serum lipids and lipoproteins from childhood to adulthood. *Am J Epidemiol.* 1991;133:884–899.
13. Nicklas TA, Farris RP, Smoak CG, et al. Dietary factors relate to cardiovascular risk factors in early life. *Arteriosclerosis.* 1988; 8:193–199.
14. Basch CE, Shea S, Contento I, Zybert P. Stability of diet measured by a semiquantitative food frequency questionnaire in 4- to 7-year-old Latino children. Unpublished manuscript.
15. Nicklas TA, Farris RP, Johnson CC, Webber LS, Berenson GS. *Food Sources of Nutrients: A Tool for Dietary Management and Health.* New Orleans, La: National Center for Cardiovascular Health, Louisiana State University Medical Center; 1990.
16. Willett WC, Stampfer MJ, Colditz GA, et al. Relation of meat, fat, and fiber intake to the risk of colon cancer in a prospective study among women. *N Engl J Med.* 1990; 323:1664–1672.
17. Jenkins DJA, Wolever TMS, Vuksan V, et al. Nibbling versus gorging: metabolic advantages of increased meal frequency. *N Engl J Med.* 1989;321:929–934.
18. Willett WC. Challenges for public health nutrition in the 1990s. *Am J Public Health.* 1990;80:1295–1297.
19. Zuckerman AE. *Dietary Recall Nutrient Analysis Software.* Washington, DC: Alan E. Zuckerman; 1987.
20. Adams CF, Richardson M. *Nutritive Value of Foods.* Washington, DC: US Govt Printing Office; 1981. US Dept of Agriculture, Home and Garden Bulletin 72.
21. Gebhardt SE, Matthews RH. *Nutritive Value of Foods.* Washington, DC: US Govt Printing Office; 1986. US Dept of Agriculture, Home and Garden Bulletin 72.
22. De Reguero LC, de Santiago SMR. *Tabla de composicion de alimentos de uso corriente en Puerto Rico.* Hato Rey, Puerto Rico: Universidad de Puerto Rico; 1978.
23. Basch CE, Shea S, Arliss A, et al. Validation of mothers' reports of dietary intake by four to seven year-old children. *Am J Public Health.* 1990;80:1314–1317.
24. Eck LH, Klesges RC, Hanson CL. Recall of a child's intake from one meal: are parents accurate? *J Am Diet Assoc.* 1989;89: 784–789.
25. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol.* 1985;122:51–65.
26. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analysis. *Am J Epidemiol.* 1986;124:17–27.
27. Committee on Nutrition, American Acad-

emy of Pediatrics. Prudent lifestyle for children: dietary fat and cholesterol. *Pediatrics.* 1986;78:521–525.
28. Pugliese M, Weyman-Daum M, Moses N, et al. Parental health beliefs as cause of non-organic failure to thrive. *Pediatrics.* 1987;80:175–182.
29. Lifshitz F, Moses N. Growth failure: a complication of dietary treatment of hypercholesterolemia. *Am J Dis Child.* 1989; 143:537–542.
30. Sigman M, Neuman C, Jansen AAJ, Bwibo N. Cognitive abilities of Kenyan children in relation to nutrition, family characteristics, and education. *Child Dev.* 1989;60:1463–1474.
31. Sigman M, Neuman C, Baksh M, Bwibo N, McDonald A. Relationship between nutrition and development in Kenyan toddlers. *J Pediatr.* 1989;115:357–364.
32. Allen LH, Black AK, Backstrand JR, et al. An analytical approach for exploring the importance of dietary quality vs quantity to the growth of Mexican children. *Food Nutr Bull.* 1991. In press.
33. Shea S, Basch CE, Zybert P. Relationships between 4- to 7-year-old children's dietary fat intake and growth. Unpublished manuscript.
34. Gutin B, Basch CE, Shea S, et al. Blood pressure, fitness, and fatness in 5- and 6-year-old children. *JAMA.* 1990;264:1123–1127.
35. Gortmaker S, Dietz W, Sobol A, Wehler C. Increasing pediatric obesity in the United States. *Am J Dis Child.* 1987;141:535–540.
36. Hegstead DM, McGandy RB, Myers ML, Stare FJ. Quantitative effects of dietary fat on serum cholesterol in man. *Am J Clin Nutr.* 1965;17:281–295.
37. Bonanome A, Grundy SM. Effect of dietary steric acid on plasma cholesterol and lipoprotein levels. *N Engl J Med.* 1986; 318:1244–1248.
38. US Department of Commerce, Bureau of the Census. The Hispanic population in the United States: March 1988 advance report. *Current Population Reports.* Series P-20, no. 431. Washington, DC: US Govt Printing Office; 1988. USGPO publication 1988–201–478:80117.
39. Hazuda HP, Stern MP, Gaskill SP, Haffner SM, Gardner LI. Ethnic differences in health knowledge and behaviors related to the prevention and treatment of coronary heart disease: the San Antonio Heart Study. *Am J Epidemiol.* 1983;117:717–728.
40. Samet JM, Coultas DB, Howard CA, et al. Diabetes, gallbladder disease, obesity and hypertension among Hispanics in New Mexico. *Am J Epidemiol.* 1988;128:1302–1311.
41. Hamman RF, Marshall JA, Baxter J, et al. Methods and prevalence of non-insulin dependent diabetes mellitus in a biethnic population: the San Luis Valley Diabetes Study. *Am J Epidemiol.* 1989;129:295–311.
42. Mitchell BD, Stern MP, Haffner SM, Hazuda HP, Patterson JK. Risk factors for cardiovascular mortality in Mexican Americans and non-Hispanic Whites: the San Antonio Heart Study. *Am J Epidemiol.* 1990;131:423–433.
43. Stern MP, Haffner SM. Type II diabetes and its complications in Mexican Americans. *Diabetes Metab Rev.* 1990;6:29–45.
44. Shea S, Stein AD, Basch CE, et al. Independent associations of educational attainment and ethnicity with behavioral risk factors for cardiovascular disease. *Am J Epidemiol.* 1991;134:567–583.
45. Knapp, JA, Hazuda HP, Haffner SM, et al. A saturated fat/cholesterol avoidance scale: sex and ethnic differences in a biethnic population. *J Am Diet Assoc.* 1988;88:172–177.
46. Delapa RM, Mayer JA, Candelaria J, et al. Food purchase patterns in a Latino community: Project Salsa. *J Nutr Educ.* 1990; 22:133–136.
47. Miller JJ, Short SD, Neff SA. *Dairy: situation and outlook yearbook.* Rockville, Md: US Dept of Agriculture, Economic Research Service, Commodity Economics Division; 1991. Publication DS–431.
48. Wechsler H, Wernick SM. A social marketing campaign to promote the consumption of low-fat milk in an inner-city Latino community. *Public Health Rep.* In press.

APPENDIX—Thirty-three Groups Derived from a US Department of Agriculture Typology

1	Low-fat milk products (<10% of calories from fat)
2	Medium-fat milk products (10–40% of calories from fat)
3	High-fat milk products (40% of calories from fat)
4	Beef/veal
5	Pork/lamb/sausage (not smoked)
6	Poultry, not fried
7	Poultry, fried
8	Fish/shellfish, not fried
9	Fish/shellfish, fried
10	Cured and processed meats/organ meats/franks/lunchmeat/bacon/ham
11	Meat mixtures/sandwiches
12	Eggs/egg mixtures
13	Legumes/nuts/seeds
14	Enriched bread/rolls/muffins/rice/noodles/pasta/biscuits/flour
15	100% whole grain bread/rolls/muffins/rice/noodles/pasta/biscuits/flour
16	Cakes/cookies/pies/pastries/doughnuts/graham crackers
17	Crackers/chips/salty snacks from grains
18	Breakfast cereals
19	Pancakes/waffles/french toast
20	Grain mixtures, mostly grain
21	Fresh fruit
22	Other fruit
23	Fruit juice
24	Dark green leafy and nonleafy vegetables/deep yellow vegetables
25	Other vegetables/vegetable mixtures
26	Vegetables, fried
27	Fats and oils used in cooking and at the table from vegetable sources
28	Salad dressing/gravy
29	Fats and oils used in cooking and at the table from animal sources
30	Sugars/jellies/syrups/gelatin/desserts/popsicles/ices/candies/gum/puddings/icing
31	Soda/punch/fruit drinks
32	Alcoholic beverages
33	Condiments

Source. Data from the US Department of Agriculture.[21]

Socioeconomic Status and Health: How Education, Income, and Occupation Contribute to Risk Factors for Cardiovascular Disease

Marilyn A. Winkleby, PhD, Darius E. Jatulis, MS, Erica Frank, MD, MPH, and Stephen P. Fortmann, MD

ABSTRACT

Background. Socioeconomic status (SES) is usually measured by determining education, income, occupation, or a composite of these dimensions. Although education is the most commonly used measure of SES in epidemiological studies, no investigators in the United States have conducted an empirical analysis quantifying the relative impact of each separate dimension of SES on risk factors for disease.

Methods. Using data on 2380 participants from the Stanford Five-City Project (85% White, non-Hispanic), we examined the independent contribution of education, income, and occupation to a set of cardiovascular disease risk factors (cigarette smoking, systolic and diastolic blood pressure, and total and high-density lipoprotein cholesterol).

Results. The relationship between these SES measures and risk factors was strongest and most consistent for education, showing higher risk associated with lower levels of education. Using a forward selection model that allowed for inclusion of all three SES measures after adjustment for age and time of survey, education was the only measure that was significantly associated with the risk factors ($P < .05$).

Conclusion. If economics or time dictate that a single parameter of SES be chosen and if the research hypothesis does not dictate otherwise, higher education may be the best SES predictor of good health.(*Am J Public Health.* 1992;82:816–820)

Introduction

One of the strongest and most consistent predictors of a person's morbidity and mortality experience is that person's socioeconomic status (SES).[1–5] This finding persists across all diseases with few exceptions, continues throughout the entire life span,[5] and extends across numerous risk factors for disease.[6–9]

The significant impact of SES on disease makes its definition and measurement of critical importance. SES is a complex phenomenon predicted by a broad spectrum of variables that is often conceptualized as a combination of financial, occupational, and educational influences.[10–13] Although these dimensions of SES are interrelated, it has been proposed that each reflects somewhat different individual and societal forces associated with health and disease. For example, income reflects spending power, housing, diet, and medical care; occupation measures prestige, responsibility, physical activity, and work exposures; and education indicates skills requisite for acquiring positive social, psychological, and economic resources.[9,14]

Much has been written about the underlying mechanisms through which SES may operate to affect disease. Many authors have suggested that certain dimensions of SES are more predictive of health than others;[1,14–16] however, such proposals tend to be theoretically based, without substantiating data. Over time, education has become the most commonly used measure of SES in epidemiological studies,[17] yet no investigators in the United States have conducted an empirical analysis quantifying the relative contributions of different measures of SES to risk factors or disease outcomes. (This paper does not suggest which component of SES may be the most reliable and valid measure because this question has been examined, with a variety of conclusions, in previously published reports.)[10–13, 17–18]

The present study examines the association between income, education, occupation, and a set of risk factors for cardiovascular disease—namely, cigarette smoking, systolic and diastolic blood pressure, and total and high-density lipoprotein (HDL) cholesterol. Our study objectives are to (1) examine the impact of each separate dimension of SES on cardiovascular risk factors, (2) use a forward stepwise selection model to evaluate if one measure of SES is the strongest predictor of risk factors, and (3) offer guidance to researchers about selection of SES measures. This guidance is critical because nearly all epidemiological studies use SES as an explanatory or a control variable, or for the selection of subjects or matching criteria.

Methods

Subjects aged 25 to 64 were drawn from the two control cities of the Stanford Five-City Project,[19] a communitywide cardiovascular disease intervention study that contains data from four separate

The authors are with the Stanford Center for Research in Disease Prevention, Stanford University School of Medicine, in Palo Alto, CA.

Requests for reprints should be sent to Marilyn A. Winkleby, Stanford Center for Research in Disease Prevention, Stanford University School of Medicine, 1000 Welch Road, Palo Alto, CA 94304-1885.

This paper was submitted to the Journal April 19, 1991, and accepted with revisions September 13, 1991.

Editor's Note. See related editorial on p 785 of this issue.

cross-sectional surveys, conducted from 1979 to 1986. Participants who were unemployed (n = 98), students (n = 130), or retirees (n = 146) were excluded because they had no occupation that could be ranked. Data from the four cross-sectional surveys were pooled because demographic characteristics, including measures of SES, showed no significant interactions over time.[7] As previously reported,[20] the educational attainment of respondents from the control cities were well matched to those from the treatment cities. A nonrespondent questionnaire indicated that participants were generally more educated than those in both treatment and control communities who refused to participate.[20]

Information on education, income, and occupation was ascertained through questionnaires. Education was recorded as the highest number of years of schooling completed. It was used as a continuous variable in regression analyses and was divided into the following four categories in stratified analyses: less than 12 years, 12 years, 13 to 15 years, and 16 years or more.

Income information, defined as gross annual household income, was collected in intervals of $5000. For multivariate analyses, the midpoint of the income categories was used; for the remaining analyses, income was divided into five categories ranging from less than $10 000 to $40 000 or more per year.

Current occupation was collected as an open-ended variable and coded using the 1980 US Bureau of the Census occupational scaling system, which creates a hierarchy of occupations ordered on the basis of education and income.[21] To examine associations between separate occupations and risk factors, we coded each occupation, including homemaker, as an indicator variable, with executives as the reference category. In multivariate models, we excluded homemakers and used occupation as a ranked variable ranging from executives to unskilled workers.

Participants were considered cigarette smokers if they reported ever smoking on a daily basis and had smoked one or more cigarettes in the last week. Participants whose plasma thiocyanate exceeded 100 μmol/L and whose expired-air carbon monoxide level exceeded 8 ppm were classified as smokers regardless of their self-reported responses.[22] Total and HDL plasma cholesterol were derived from nonfasting venous samples, analyzed fresh by standard methods established by the Lipid Research Clinics Program.[23] Three blood pressure measurements were taken on the right arm using a semiautomatic recorder, and the average of the second and third readings were used for analyses.[24]

Results

An overall response rate of 69% was achieved. Approximately 600 individuals participated in each survey, resulting in a total of 2380 participants for all surveys combined. Because the study population was predominantly White, non-Hispanic (85%), findings are not generalizable to populations representing a broad spectrum of racial groups.

All pairwise correlations between education, income, and occupation were positive and were stronger for men than for women (Table 1). The lowest correlation was between education and income, indicating that education is not a primary determinant of wage. Higher correlations were shown for education and occupation, suggesting that skills acquired during education may help determine occupation. Although correlations ranged from .23 to .67, their relatively low magnitude (highest adjusted R^2 = 45%) indicates that the three dimensions are not redundant measures of SES.

Although the study population had relatively high educational and financial levels, individuals from all education, income, and occupation categories were well-represented (Table 2). However, while men were represented fairly evenly across occupations, approximately 75% of women employed outside the home held nonprofessional white-collar jobs. Men tended to have more years of education than women and to be from higher-income households.

TABLE 1—Correlation Coefficients for Education, Income, and Occupation

	Men	Women
Education/income	+.32	+.23
Income/occupation	+.41	+.30
Education/occupation	+.67	+.66

Note. Table excludes homemakers. All correlations are significant at $P \leq .0001$.

In general, those with the lowest educational attainment exhibited the highest prevalence of risk factors (Table 3). Clear gradients were seen between educational level and smoking for both sexes, and between education and total and HDL cholesterol for women. Across all risk factors, men consistently exhibited higher risk than did women.

Income and occupation were less consistent risk predictors. Higher risk was associated with lower incomes for smoking and HDL cholesterol in both sexes but with higher incomes for total cholesterol in men. Within occupations, men and women white-collar executives and managers exhibited the lowest levels of smoking. Among men, executives and managers showed the lowest mean levels of blood pressure.

TABLE 2—Distribution of Study Population by Education, Income, and Occupation

	Men, % (n = 1108)	Women, % (n = 1272)
Education, y		
<12	9.7	13.1
12	23.9	31.8
13–15	24.9	27.8
≥16	41.5	27.4
Household income		
<$10 000	6.2	13.3
$10 000–$19 999	21.1	25.6
$20 000–$29 999	26.6	25.6
$30 000–$39 999	20.6	17.8
≥$40 000	25.5	17.8
Occupation		
Unskilled	6.1	5.4
Semiskilled	10.7	6.8
Skilled	17.3	3.5
Clerical	13.4	25.1
Administrators	19.9	9.3
Managers	17.8	12.7
Executives	14.7	1.7
Homemakers	0.0	35.7

Note. Table includes homemakers, all of whom were women.

TABLE 3—Prevalence of Risk Factors by Education, Income, and Occupation, Adjusted for Age and Time of Survey

	% Cigarette Smokers		Mean Systolic Blood Pressure		Mean Diastolic Blood Pressure		Mean Total Cholesterol		Mean HDL Cholesterol	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Education, y										
<12	47	41	126	117	76	73	198	201	44	52
12	41	34	127	118	77	74	198	194	44	55
13–15	31	23	126	116	77	73	203	192	44	58
≥16	18	14	125	114	76	71	196	190	46	61
Household income										
<$10 000	50	32	124	116	77	72	190	192	43	54
$10 000–$19 999	30	31	126	117	76	72	193	197	45	57
$20 000–$29 999	31	25	126	115	75	73	200	195	45	55
$30 000–$39 999	26	24	126	118	76	74	200	193	46	59
≥$40 000	25	19	126	116	77	72	202	187	46	59
Occupation										
Unskilled	36	33	128	115	78	72	191	199	46	50
Semiskilled	39	22	126	117	76	74	198	187	44	56
Skilled	37	44	127	117	77	71	195	195	46	58
Clerical	41	31	126	117	77	74	206	194	42	57
Administrators	27	25	126	115	77	72	201	190	44	58
Managers	21	18	125	114	76	70	195	190	45	61
Executives	16	20	124	115	75	74	200	192	48	58
Homemakers	n/a	25	n/a	117	n/a	73	n/a	195	n/a	56

*Note.*Tables includes homemakers. A logistic regression model was used for smoking; a multiple regression model was used for blood pressure, total cholesterol, and HDL cholesterol. Lower mean levels of HDL cholesterol indicate higher risk.

TABLE 4—Partial Correlations and *P* Values from Univariate and Forward Selection Regression Models for Three Dimensions of SES and Disease Risk Factors

	Cigarette smoking		Systolic Blood Pressure		Diastolic Blood Pressure		Total Cholesterol		HDL Cholesterol	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Univariate Model[a]										
Education	−.23***	−.18***	−.05	−.08*	−.06	−.12***	−.02	−.07*	+.11***	+.21***
Income	−.08**	−.11**	−.00	−.03	+.02	+.02	+.05	−.04	+.06*	+.08*
Occupation	−.18***	−.08*	−.06	−.05	+.05	+.06	+.02	−.03	+.04	+.16***
Forward Selection Model[b]										
Education	−.23***	−.18***	...	−.08*	...	−.12***	...	−.07*	+.11***	+.21***
Income	...	...	...	...	...	...	...	...	...	...
Occupation	...	...	...	...	...	...	...	...	...	...

*Note.*Table excludes homemakers. Two-tailed significance values are coded as $^{*}P < .05$, $^{**}P < .01$, and $^{***}P < .001$.
[a]Adjusted for age and time of survey.
[b]Model allowed for inclusion of education, income, and occupation after adjusting for age and time of survey. Ellipses indicate not selected for inclusion in model at $P < .05$.

To compare the strength of the interrelationships between the three dimensions of SES, we conducted regression analyses (Table 4), adjusted for age and time of survey. Partial correlations are presented with two-tailed *P* values, with significance defined as $P < .05$. The univariate relationship between SES and risk factors was strongest and most consistent for education, showing higher risk associated with lower levels of education. Using a forward selection model that allowed for inclusion of all three measures of SES after adjustment for age and time of survey, education was the only measure of SES that was significantly associated with the risk factors. This finding is consistent with a previously conducted multivariate analysis, which showed that the strength and significance of the associations between education and a set of disease risk factors remained virtually unchanged after adjustment for income and occupation.[7]

When stratified by White, non-Hispanic and Hispanic ethnicities (not shown), education remained the strongest SES predictor but became nonsignificant for smoking, possibly because of the small sample of Hispanics.

Discussion

Strengths and Limitations of Using Education as the Marker for SES

We caution that, in some studies, using only one indicator of SES may yield misleading results or provide less information than using multiple measures. However, using multiple or composite measures[10–14] requires the cost and time of collecting data on several SES parameters and may not significantly explain more about a population than would a single, well-chosen parameter. As noted by a re-

cent working group of the National Heart, Lung, and Blood Institute, use of composite measures may obscure important differences in associations.[25]

Based on our findings, education may be the most judicious SES measure for use in epidemiological studies (unless the study hypothesis dictates which dimension of SES is to be chosen). In studies that have a cost or time restraint but need a measure of SES as a potential confounding variable, education is an expeditious choice. In addition, education is available for all individuals regardless of employment status, has high reliability and validity,[17] is generally stable after early adulthood, is easily reported, and can be collected as a continuous variable. Furthermore, because education is often available in epidemiological studies,[17] it also permits opportunities for meta-analyses and interstudy comparisons.

However, there are potential limitations to using education as a sole indicator of SES. Its stability can mask important changes in individuals' circumstances. There may also be a cohort effect distorting differences between populations of various ages. For example, the percent of the population obtaining at least a high school education has increased nearly threefold since 1940.[17] This has led to increasing homogeneity in the amount of education obtained, making differentiation between educational strata more difficult. Other potential problems include regional differences in education, the question of whether degrees or certification are better measurement parameters than years of schooling,[17,26,27] and the possibility that other dimensions of SES are more sensitive markers for health in some population subgroups. (For example, for foreign-born female Hispanics, acculturation may be a stronger measure of SES.)[13]

Comparison with Past Studies

To our knowledge, this is the only study in the United States to examine associations between separate SES dimensions and risk factors or disease outcomes. Other studies, however, have examined associations between one measure of SES and one disease risk factor,[6,8,28–32] morbidity,[33] or mortality.[1,33–35] In general, these studies have found that education is more strongly associated with disease than is income or occupation. One of the most complete studies of mortality differentials found that lower SES groups exhibited higher rates of all-cause mortality than did higher SES groups, irrespective of whether education, income, or occupation was used as the measure of SES.[1] Other studies have documented strong inverse associations between education and all-cause mortality[31,34,36] and life expectancy.[37] Framingham study data shows that, of 23 potential contributors to morbidity, only education and age at study enrollment were related in both sexes to "survival with good function."[38] Cardiovascular disease studies have shown that lower levels of education are associated with hypertension,[6,28–29,31,39] cigarette smoking,[6,31,32,40–41] and high cholesterol,[6,32] as well as with cardiovascular morbidity[33] and mortality.[31,33,35]

Why Education May Be the Strongest Predictor of Good Health

Several different mechanisms through which education may positively influence health have been proposed. Fuchs has suggested that both education and health are markers for willingness to delay gratification in order to "invest in human capital."[15] Others have argued that education may simply serve as a marker for intelligence; however, studies showing that environmental factors are the strongest predictors of school dropout lend little support to this hypothesis.[42] Some have suggested that higher education may improve health by conferring economic advantages, but our low correlation coefficients between education and income, as well as the lack of significant income effects on risk factors, argue against this hypothesis. Neither does the health knowledge acquisition that accompanies higher education appear to explain the relationship between education and health, given that provision of information alone appears to be a weak stimulus to human behavior change.[16,37]

One hypothesis we find most plausible is that education may protect against disease by influencing life-style behaviors, problem-solving abilities, and values.[17] Moreover, education may facilitate the acquisition of positive social, psychological, and economic skills and assets, and may provide insulation from adverse influences.[7] Such skills and assets that may accompany higher educational attainment include positive attitudes about health, access to preventive health services,[15] membership in peer groups that promote the adoption or continuation of positive health behaviors, and higher self-esteem and self-efficacy.[43,44]

Summary

There can be no SES measure that is universally valid and suitable for all populations. However, if economics and time dictate that a single parameter be chosen and if the research hypothesis does not dictate otherwise, this study suggests that higher education, rather than income or occupation, may be the strongest and most consistent predictor of good health. □

Acknowledgments

This research was supported by Public Health Service Grant 1RO1–HL–21906 from the National Heart, Lung, and Blood Institute to Dr. John W. Farquhar.

The authors thank Drs. Lawrence Green and David Ragland for comments on an earlier draft, Dr. Helena Kraemer and Ms. Beverly Rockhill for statistical advice, and Ms. Cindy German-Kung for the preparation of tables.

References

1. Kitagawa EM, Hauser PM. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology.* Cambridge, Mass: Harvard University Press; 1973.
2. Blaxter M. Evidence on inequality in health from a national survey. *Lancet* 1987;ii:30–33.
3. Black D. *Inequalities in Health: Report of a Research Working Group.* London, England: Department of Health and Social Security; 1980.
4. Haan M, Kaplan G, Camacho T. Poverty and health: prospective evidence from the Alameda County Study. *Am J Epidemiol.* 1987;125:989–998.
5. Marmot MG, Kogevinas M, Elston MA. Social/economic status and disease. *Ann Rev Public Health.* 1987;8:111–135.
6. Matthews KA, Kelsey SF, Meilahn EN, Kuller LH, Wing RR. Educational attainment and behavioral and biological risk factors for coronary heart disease in middle-aged women. *Am J Epidemiol.* 1989;129:1132–1144.
7. Winkleby MA, Fortmann SP, Barrett DC. Social class disparities in risk factors for disease: eight-year prevalence patterns by level of education. *Prev Med.* 1990;19:1–12.
8. Helmert U, Herman B, Joeckel KH, Greiser E, Madans J. Social class and risk factors for coronary heart disease in the Federal Republic of Germany: results of the baseline survey of the German Cardiovascular Prevention Study. *J Epidemiol Community Health.* 1989;43:37–42.
9. Antonovsky A. Social class, life expectancy and overall mortality. *Milbank Mem Fund Q.* 1967;45:31–73.
10. Mueller CW, Parcel TL. Measures of socioeconomic status: alternatives and recommendations. *Child Dev.* 1981;52:13–30.
11. Hollingshead AB. The indiscriminate state of social class measurement. *Soc Forces.* 1971;49:563–567.
12. Duncan OD. A socioeconomic index for all occupations. In: Reiss AJ Jr, Duncan OD, Hatt PK, North CC, eds. *Occupation and Social Status.* New York: The Free Press of Glencoe; 1961:109–138.
13. Green L. Manual for scoring socioeconomic status for research on health behavior. *Public Health Rep.* 1970;85:815–827.
14. Susser MW, Watson W, Hopper K. *Soci-*

ology in Medicine. New York, NY: Oxford University Press; 1985.
15. Fuchs VR. Economics, health, and post-industrial society. *Milbank Mem Fund Q/Health and Society.* 1979;57:153–182.
16. Williams DR. Socioeconomic differentials in health: a review and redirection. *Soc Psych Q.* 1990;53:81–99.
17. Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev.* 1988;10:87–121.
18. Morgenstern H. Socioeconomic factors: concepts, measurement and health effects. NHLBI Workshop on "Measuring Psychosocial Variables in Epidemiological Studies of Cardiovascular Disease." December 1983; Galveston, TX.
19. Farquhar JW, Fortmann SP, Maccoby N, et al. The Stanford Five-City Project: design and methods. *Am J Epidemiol.* 1985; 122:323–334.
20. Farquhar JW, Fortmann SP, Flora JA, et al. Effects of communitywide education on cardiovascular disease risk factors. *JAMA.* 1990;264:359–365.
21. US Bureau of the Census. *1980 Census of Population: Alphabetical Index of Industries and Occupations.* Washington, DC: US Government Printing Office; 1982.
22. Fortmann SP, Rogers T, Haskell WL, Solomon DS, Vranizan K, Farquhar JW. Indirect measures of cigarette use: expired air carbon monoxide vs plasma thiocyanate. *Prev Med.* 1984;13:127–135.
23. US Department of Health, Education, and Welfare. *Lipid Research Clinics Manual of Laboratory Operations.* Vol 1: *Lipid and Lipoprotein Analysis.* Washington, DC: US Government Printing Office; 1974.
24. Fortmann SP, Marcuson R, Bitter PH, Haskell WL. A comparison of the Sphygmetrics SR-2 automatic blood pressure recorder to the mercury sphygmomanometer in population studies. *Am J Epidemiol.* 1981; 114:836–844.
25. Socioeconomic factors: working group summary. In: Ostfeld AM, Eaker ED, eds. *Measuring Psychosocial Variables in Epidemiologic Studies of Cardiovascular Disease: Proceedings of a Workshop.* Washington, DC: US Government Printing Office; March 1985. NIH publication 85–2270.
26. Morgan M. Measuring social inequality: occupational classifications and their alternatives. *Community Med.* 1983;5:116–124.
27. Faia MA. Selection by certification: a neglected variable in stratification research. *Am J Sociol.* 1981;86:1093–1111.
28. Hypertension Detection and Follow-Up Program Cooperative Group. Race, education and prevalence of hypertension. *Am J Epidemiol.* 1977;106:351–361.
29. Dyer AR, Stamler J, Shekelle RB, Schoenberger J. The relationship of education to blood pressure: findings on 40 000 employed Chicagoans. *Circulation.* 1976;54: 987–992.
30. Abramson JH, Gofin R, Habib J, Pridan H, Gofin J. A comparative appraisal of measures for use in epidemiological studies. *Soc Sci Med.* 1982;16:1739–1746.
31. Liu K, Cedres LB, Stamler J, et al. Relationship of education to major risk factors and death from coronary heart disease, cardiovascular diseases, and all causes. *Circulation.* 1982;66:1308–1314.
32. Jacobsen BK, Thelle DS. Risk factors for coronary heart disease and level of education. *Am J Epidemiol.* 1988;127:923–932.
33. Hinkle LE, Whitney LH, Lehman EW, et al. Occupation, education, and coronary heart disease. *Science.* 1968;161:238–246.
34. Snowdon DA, Ostwald SK, Kane RL. Education, survival and independence in elderly Catholic sisters, 1936–1988. *Am J Epidemiol.* 1989;130:999–1012.
35. Hypertension Detection and Follow-Up Program Cooperative Group. Educational level and 5-year all-cause mortality in the Hypertension Detection and Follow-Up Program. *Hypertension.* 1987;9:641–646.
36. Feldman JJ, Makuc DM, Kleinman JC, Cornoni-Huntley J. National trends in educational differentials in mortality. *Am J Epidemiol.* 1989;129:919–933.
37. Sagan LA. *The Health of Nations.* New York, NY: Basic Books, Inc; 1987.
38. Pinsky JL, Leaverton PE, Stokes J. Predictors of good function: the Framingham study. *J Chronic Dis.* 1987;40:159S–167S.
39. Berger MC, Leigh JP. Schooling, self-selection, and health. *J Hum Res.* 1989;24: 433–455.
40. Millar WJ, Wigle DT. Socioeconomic disparities in risk factors for cardiovascular disease. *Can Med Assoc J.* 1986;134:127–132.
41. Wagenknecht LE, Perkins LL, Cutler GR, et al. Cigarette smoking is strongly related to educational status: the CARDIA study. *Prev Med.* 1990;19:158–169.
42. Howard MA, Anderson RJ. Early identification of potential school dropouts: a literature review. *Child Welfare.* 1978;57:221–231.
43. Bunker JP, Gomby DS. Preface. In: Bunker JP, Gomby DS, Kehrer BH, ed. *Pathways to Health: The Role of Social Factors.* Menlo Park, Calif: Henry J. Kaiser Foundation; 1989:15.
44. Cohen S, Syme SL. *Social Support and Health.* New York, NY: Academic Press; 1985.

Socioeconomic and Behavioral Correlates of Body Mass Index in Black Adults: The Pitt County Study

Janet B. Croft, PhD, David S. Strogatz, PhD, Sherman A. James, PhD, Nora L. Keenan, MPH, Alice S. Ammerman, DrPH, Ann M. Malarcher, PhD, and Pamela S. Haines, DrPH

ABSTRACT

Background. Obesity is more prevalent among Black women than Black men, but there is little information on the correlates of obesity in Blacks. This study describes the relations of sociodemographic factors and health behaviors to body mass index in a southern, Black population.

Methods. In 1988, a community probability sample of 1784 Black adults, aged 25 to 50, was examined in Pitt County, NC.

Results. More women than men were at least 20% overweight (57% vs 36%). The relation of socioeconomic status (a composite of education and occupation) to age-adjusted body mass index level was inverse in women but not in men. Body mass index did not differ with either current energy intake or energy expenditure. Smokers and drinkers had lower age-adjusted levels than nonsmokers and abstainers.

Conclusions. Since the excess body mass index levels associated with low socioeconomic status in women could not be explained after controlling for adverse health behaviors, further epidemiologic study of risk factors for obesity in Black women is recommended. (*Am J Public Health.* 1992;82:821–826)

Introduction

A recent National Institutes of Health consensus development conference emphasized the need for epidemiologic research on the distribution and etiology of obesity in minority and lower socioeconomic populations.[1,2] Obesity, or excess adiposity, is a major health problem in the United States, with over 34 million overweight adults (25.7%) defined as at least 20% over their desirable weight.[3] Among Black women, the prevalence is much higher, with 43.8% overweight and 19.3% severely overweight (at least 40% over their desirable weight).[3] Considering that Black women have twice the overweight prevalence of White women, it is not surprising that Black–White comparisons of US mortality rates and morbidity prevalence suggest an excess of obesity-related health outcomes among Black women.[4] This excess calls for more extensive investigations of potential environmental and behavioral risk factors of obesity in this group.

Socioeconomic status (SES) is an important correlate of obesity level in US populations.[5–9] Body mass index (BMI) level, skinfold measurements, and overweight prevalence in Black women have consistent inverse relationships to SES indicators, including educational level, poverty status, and income level.[7,10] A weaker positive relationship between SES and BMI has been observed in Black men.[8–11] SES probably reflects broad environmental exposures as well as specific behavior patterns. The association of low SES with increased use of alcohol and cigarette smoking,[12,13] increased unemployment,[13] and a higher caloric intake[14] has been observed in the economically heterogeneous adult Black population of Pitt County, NC. Literature reviews suggest inverse relationships of both cigarette smoking and alcohol use with obesity in men and women.[4,15,16] Although obesity is assumed to be a consequence of the chronic imbalance between energy intake and expenditure,[17] a review of obesity in Black women suggested that excess obesity may not be related to excess caloric intake.[4] This study provides an opportunity to further examine the relationship of these health behaviors to BMI, a surrogate measure of obesity, and to examine whether the association of SES with BMI can be explained by these life-style behaviors.

Janet B. Croft, David S. Strogatz, Sherman A. James, and Nora L. Keenan are with the Department of Epidemiology, and Alice S. Ammerman and Pamela S. Haines are with the Department of Nutrition, School of Public Health, at the University of North Carolina at Chapel Hill. David S. Strogatz is also with the Department of Epidemiology, School of Public Health, State University of New York at Albany and the New York State Department of Health, Albany. Sherman A. James is also with the Department of Epidemiology, School of Public Health, and Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor. Ann M. Malarcher is with the Research Institute on Alcoholism, Buffalo, NY.

Requests for reprints should be sent to Sherman A. James, PhD, University of Michigan, Department of Epidemiology, School of Public Health, 109 Observatory Street, Ann Arbor, MI 48109.

This paper was submitted to the Journal April 9, 1991, and accepted with revisions September 6, 1991.

Editor's Note. See related editorial on p 785 of this issue.

Methods

The overall purpose of the Pitt County Study is to determine psychosocial and dietary risk factors for elevated blood pressure in an eastern North Carolina population of young and early-middle-aged Black adults. In 1988, a stratified random sample of all dwellings within Pitt County (33% Black, 66% White) was used to select a community probability sample of 25- to 50-year-old Black adults. Selection of the probability sample has been described elsewhere.[12] Among the 2225 age-eligible Black residents, interviews were obtained throughout the year in the homes of 1784 (80%) adults.

After the respondents removed shoes and heavy clothing, trained interviewers obtained measurements of weight (in pounds) with a Healthometer portable scale and height (to the nearest inch) with a vertical ruler. BMI was computed as weight in kilograms divided by a squared term of height in meters. Three BMI categories were created using the 1983 Metropolitan Life Insurance Company relative weight tables[2,10,18,19] as a standard. There were no reliability assessments of height and weight measurements.

A standardized questionnaire was completed by each participant during the interview and included questions on health habits, education, occupation, and food frequency. Detailed descriptions of questionnaire components are provided elsewhere.[12–14,20] The food frequency component[21] was modified to reflect regional food practices and was administered as part of the interview to obtain estimates of daily intakes of nutrients and total calories (an indicator of energy intake).[14] Alcohol consumption was divided into four categories, ranging from abstention to seven or more drinks per week.[12] Other health habits included current cigarette smoking and physical activity. Strenuous physical activity (a general indicator of energy expenditure) was based on self-reports of sweat-producing work or exercise lasting for 20 minutes or more, at least three times a week.[20,22] Participants were defined as married if they reported either being married or living with a companion. Individuals who had previously worked but were not working at the time of the interview were characterized as currently unemployed; currently employed persons, homemakers, students, and individuals who had never worked were characterized as "other." An SES index combining education and occupational status was developed.[13,14] The low SES category included unskilled workers who had a high school education or less and semiskilled workers who had not completed high school. The intermediate SES category included unskilled workers who had education beyond the high school level, semiskilled workers who had at least completed high school, and participants who had not finished high school but reported skilled craftsman or white-collar occupations. High SES classified those who had a high school education or greater and reported skilled craftsman or white-collar occupations.

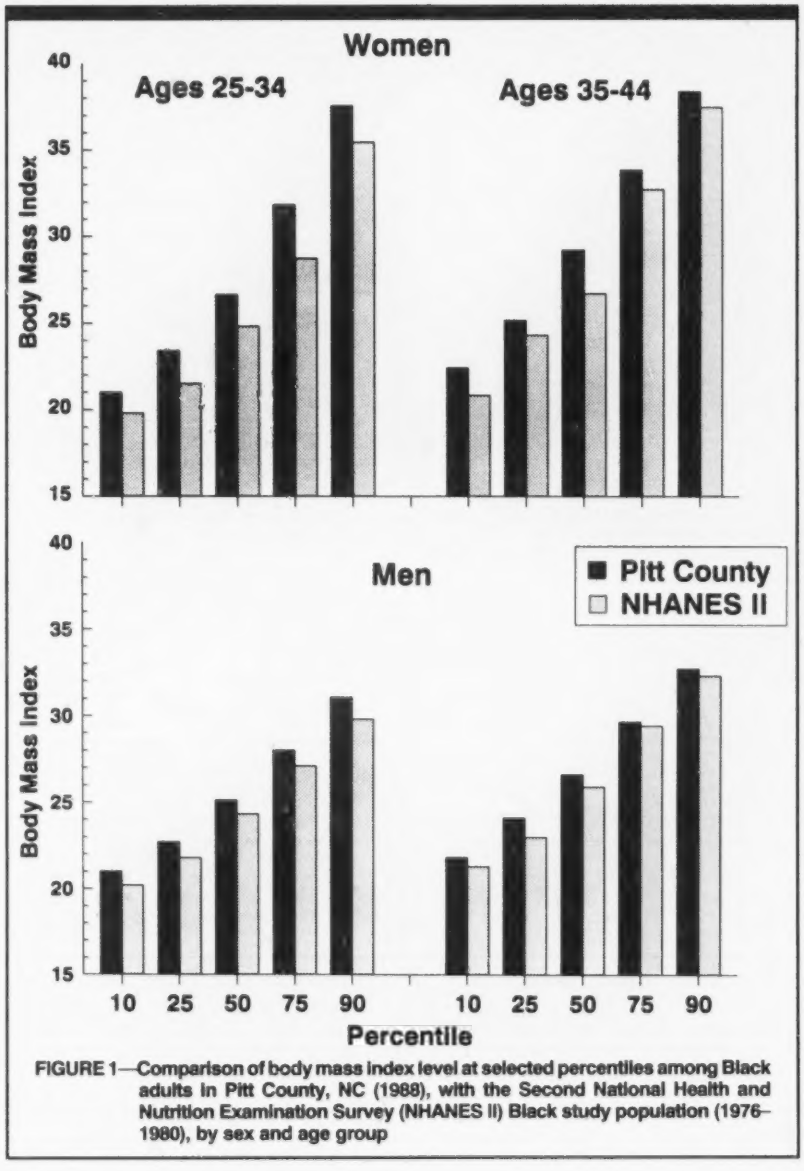


FIGURE 1—Comparison of body mass index level at selected percentiles among Black adults in Pitt County, NC (1988), with the Second National Health and Nutrition Examination Survey (NHANES II) Black study population (1976–1980), by sex and age group

Individuals missing values for height, weight, relevant health behaviors, or SES were not included in the final analyses, resulting in a total of 1663 (93%) individuals with complete data. All analyses were sex specific and weighted to take into account the sampling strategy as well as nonresponse. Sex-specific distributions of categories of BMI, sociodemographic characteristics, and health behaviors were obtained to describe the study population. Age-adjusted mean BMI and the difference in mean BMI level between categories of each potential correlate were obtained with age-adjusted linear regression techniques. The 95% confidence interval for each difference was computed using standard errors obtained with SUREGR.[23] Multivariable linear regression techniques with BMI as the dependent

variable were used to assess correlates of BMI. Product terms combining smoking status and specific levels of alcohol consumption were added to the multivariable model to assess whether there was an interactive effect of alcohol and tobacco use on BMI level.[16] In a separate analysis, a product term combining exercise level and caloric intake was added to examine whether the relation of energy intake to BMI level varied with level of energy expenditure.

Results

The sample included 633 Black men and 1030 Black women. BMI levels ranged from 15.7 to 58.7 kg/m^2 (mean = 29.4) in women and from 16.99 to 46.3 kg/m^2 (mean = 26.3) in men. Figure 1 demonstrates that age-specific BMI levels in the Pitt County Black population were higher at selected percentiles than those of Blacks in the Second National Health and Nutrition Examination Survey (NHANES II).[3]

In all, 24.1% of women and 25.7% of men were classified in the 120% through 139% overweight category (Table 1). In comparison to 10.4% of men, an additional 33.3% of women were 140% or more overweight. Over 70% of the study population was under age 40. Men were more likely than women to be married and less likely to be currently unemployed. A greater proportion of women reported sedentary activity than men, while the proportion of women consuming less than 1500 calories per day was much higher than in men. In comparison to men, more women were nonsmokers and alcohol abstainers.

Mean BMI increased with age, especially in Black women (Table 2). Age-adjusted sex-specific associations of each correlate with BMI level were then examined (Table 2). Married individuals were heavier, especially among men. The mean BMI for low SES women was 1.5 kg/m^2 higher than for high SES women; however, BMI did not vary with SES level among men. Indicators of energy intake and energy expenditure were not associated with BMI level in women and men. Smokers had lower age-adjusted BMI levels than nonsmokers among both men and women. Women who consumed alcohol were 2 to 3 kg/m^2 lighter than women who abstained, but there was no gradient with amount of drinks per week. However, among men, there was a monotonic decrease in BMI across the alcohol categories.

TABLE 1—Sex-Specific Weighted Characteristics: 25- to 50-Year-Old Black Women and Men, Pitt County, NC 1988

Characteristic	Women		Men	
	n	%	n	%
Overweight[a]				
<120%	440	42.5	407	63.9
120%–139%	249	24.1	160	25.7
≥140%	341	33.3	66	10.4
Age, y				
<40	710	71.2	434	72.7
≥40	320	28.8	199	27.3
Marital status				
Unmarried	605	53.9	273	41.5
Married	425	46.1	360	58.5
Employment				
Other	838	83.5	577	92.7
Unemployed	192	16.5	56	7.3
Socioeconomic Status				
Low	437	39.7	230	35.3
Intermediate	295	29.5	205	34.8
High	298	30.8	198	30.0
Exercise				
Sedentary	658	62.3	277	40.4
Strenuous	372	37.7	356	59.5
Energy intake, kcal				
<1500	400	40.5	118	20.7
1500–1999	274	26.3	138	20.3
2000–2499	163	15.6	113	18.7
2500–2999	92	8.3	85	12.7
≥3000	101	9.2	179	27.5
Cigarette smoking				
Nonsmoker	682	68.9	337	54.2
Smoker	348	31.1	296	45.7
Alcohol drinks per week				
None	610	62.1	241	38.4
<1	179	16.6	70	12.6
1–6	178	16.3	190	28.9
≥7	63	4.9	132	20.0

[a]Women with body mass index (BMI) levels of less than 26.9 kg/m^2 were included in the <120% overweight group; those with levels between 26.9 and 31.4 kg/m^2 were placed in the 120%–139% group; and those with levels of 31.4 kg/m^2 or more were placed in the ≥140% group. Men with BMI levels of less than 27.2 kg/m^2 were placed in the <120% overweight group; those with levels between 27.2 and 31.8 kg/m^2 were included in the 120%–139% group; and those with levels of 31.8 kg/m^2 or more were included in the ≥140% group.

Multivariable linear regression models were used to initially assess the association of sociodemographic variables to BMI level (Table 3, Model 1). Among women, age and low SES were the only independent sociodemographic correlates of BMI in this model. Marital status was the only independent sociodemographic correlate of BMI among men. The addition of health behaviors to these models (Model 2) did not account for the sociodemographic predictors identified in Model 1. Smoking and alcohol consumption were independent, negative correlates of BMI level among women, while only smoking behavior was an independent, negative correlate of BMI level in men. Similar results were obtained in separate models after using a continuous alcohol term and after substituting either the occupation or education component for the SES variable. Furthermore, there was no interaction of either smoking and alcohol consumption or physical activity and total caloric intake with BMI level in the multivariable models containing sociodemographic factors and health behaviors.

Discussion

The slightly higher BMI levels in the Pitt County population relative to NHANES II[3] could reflect a secular increase in levels since the 1976 to 1980 national survey. Secular changes in this direction have been described for young Black women between 1960 and 1980.[24] Alternatively, these data may reflect geographic differences in BMI, with more prevalent obesity and higher weight levels in the southeastern United States than in

TABLE 2—Age-Adjusted Relation of Selected Characteristics to Body Mass Index: 25- to 50-Year-Old Black Women and Men, Pitt County, NC, 1988

Characteristic	Women (n = 1030) Mean	Difference	95% Confidence Interval	Men (n = 633) Mean	Difference	95% Confidence Interval
Age, y						
<40[a]	29.0	...	...	26.2	...	...
≥40	30.4	1.4	0.4, 2.3	26.7	0.5	−0.3, 1.3
Marital status						
Unmarried[a]	29.2	...	...	25.5	...	...
Married	29.7	0.5	−0.5, 1.4	27.0	1.5	0.7, 2.2
Employment						
Unemployed[a]	29.3	...	...	26.5	...	...
Other	29.9	0.6	−0.6, 1.8	25.5	−1.0	−2.1, 0.0
Socioeconomic status						
Low	30.2	1.5	0.1, 3.0	26.2	−0.7	−1.5, 0.1
Intermediate	29.2	0.5	−0.8, 1.9	26.2	−0.7	−1.4, 0.1
High[a]	28.7	...	...	26.9	...	...
Exercise						
Sedentary[a]	29.7	...	...	26.3	...	...
Strenuous	29.0	−0.7	−1.7, 0.4	26.4	0.1	−0.8, 1.0
Energy intake, kcal						
<1500[a]	29.5	...	...	26.5	...	...
1500–1999	29.1	−0.4	−1.5, 0.7	26.5	0.0	−1.2, 1.1
2000–2499	29.7	0.2	−1.4, 1.8	26.6	0.1	−1.1, 1.3
2500–2999	29.5	0.0	−2.0, 2.0	26.4	−0.1	−2.0, 1.8
≥3000	29.3	−0.2	−1.9, 1.4	25.9	−0.6	−1.7, 0.5
Cigarette smoking						
Nonsmoker[a]	30.1	...	...	27.1	...	...
Smoker	28.0	−2.1	−3.1, −1.0	25.5	−1.6	−2.1, −1.0
Alcohol drinks per week						
None[a]	30.3	...	...	27.0	...	...
<1	28.4	−1.9	−2.9, −0.9	26.9	−0.1	−1.2, 1.0
1–6	27.4	−2.9	−4.1, −1.8	25.9	−1.1	−1.9, −0.2
≥7	28.1	−2.2	−4.0, −0.5	25.6	−1.4	−2.4, −0.3

[a]Referent.

TABLE 3—Weighted Multivariate Linear Regression Coefficients for Selected Correlates of Body Mass Index: 25- to 50-Year-Old Black Women and Men, Pitt County, NC, 1988

Correlate	Women (n = 1030) Model 1[a] β	SE	Model 2[b] β	SE	Men (n = 633) Model 1[a] β	SE	Model 2[b] β	SE
Age, y	0.10	.03*	0.09	.04*	0.04	.03	0.03	.03
Married vs unmarried	0.75	.44	0.35	.42	1.34	.38*	1.03	.40*
Unemployed vs other	0.38	.62	0.67	.53	−0.68	.58	−0.53	.61
Low vs high SES	1.60	.77*	1.90	.76*	−0.36	.40	−0.11	.44
Medium vs high SES	0.57	.68	0.61	.71	−0.58	.37	−0.42	.41
Strenuous vs sedentary exercise			−0.71	.50			0.10	.44
Energy intake, kcal			−0.0001	.0003			−0.0001	.0002
Smoker vs nonsmoker			−1.34	.55*			−1.15	.35*
<1 drink per week vs none			−1.44	.51*			0.17	.62
1–6 drinks per week vs none			−2.66	.65*			−0.52	.43
≥7 drinks per week vs none			−1.88	1.05			−0.46	.63

Note. SES = socioeconomic status.
[a]Model 1: age + marital + unemployment + Low SES + medium SES. R^2 = .03 for women and .04 for men.
[b]Model 2: Model 1 + exercise + calories + smoking + 3 alcohol terms. R^2 = .07 (both women and men).
*$P < .05$.

other regions of the country.[25,26] Geographic differences in obesity prevalence would be consistent with regional differences in hypertension and hypertension-related mortality.[27,28]

The distributions of BMI, SES, and health behaviors, as well as the relation of sociodemographic and behavioral factors with BMI, differed between men and women in this economically heterogeneous sample of Black adults. Age was more strongly associated with BMI for women than for men. Childbearing may account for some of this difference in effect, since European studies suggest that some of the weight gained during pregnancy may be retained.[29,30] The enhanced association between age and BMI in women may reflect a cumulative impact of pregnancies.[31] Information on reproductive history was not available to examine this possibility.

Low SES was related to higher BMI levels in women but not in men. These relationships are consistent with those observed for surrogate measures of obesity in national surveys[3,7,8,10] and other cross-sectional surveys of Black populations.[9] Married men, but not women, were heavier than those who were unmarried. A predominantly White adult study reported a statistically significant and somewhat curvilinear relationship between marital status and BMI in both men and women, such that the lowest and highest BMI quintiles were less likely to be married.[32] Women at higher SES levels may be disproportionately more likely to be concerned with thinness and diet,[6] while there may be less social pressure for men to conform to an ideal body image. Furthermore, in the NHANES I follow-up cohort, the women at greatest risk of weight gain were those with low family income and less than a college education and those entering marriage.[33]

Ironically, more adverse health behaviors, cigarette and alcohol use, were negatively correlated with BMI in this study population and in others.[16] The finding that cigarette smokers had significantly lower BMI levels than nonsmokers is consistent with NHANES II data in which both male and female cigarette smokers were leaner than nonsmokers.[34] Several studies have reported a stronger inverse relation of alcohol with BMI level in women than men.[35–37] Differences between men and women in the relation of alcohol intake to BMI level might be explained, in part, by the decreased metabolism and gastric oxidation of ethanol in women.[38]

As in several surveys,[4,34,39,40] the present results demonstrated that on a population basis, excess body weight was not necessarily associated with higher caloric intake. A major problem of cross-sectional studies is that self-reported dietary intake may not reflect long-term dietary patterns or take into account individuals on weight-reduction diets. In the present study population, 8% of all men reported that they were currently dieting, regardless of overweight status; however, reports of dieting increased with overweight status among women (8%, 10%, and 17% in the < 120%, 120% through 139%, and ≥ 140% overweight categories, respectively). Multivariable linear regression results were similar after controlling for current dieting.

Although over 60% of women and 40% of men reported a sedentary lifestyle, sedentary behavior was also not related to BMI level in the Pitt County population. These results are not consistent with some studies that have observed an inverse relationship between physical activity and obesity level, especially in men.[35,39,41] The difference in results may be due to methodologic factors (e.g., the measurement and definition of physical activity, seasonal variations in sweat-producing activities, the sampling strategy, and criteria for inclusion in the study). Furthermore, cross-sectional surveys often do not measure long-term physical activity patterns and are unable to determine the temporal relationship of current physical activity with obesity level.

The results of this study confirm the empirical evidence that obesity is a prevalent problem among low SES populations and Black women. Low SES was an important correlate of high BMI in Black women, but elevated levels could not be explained by adverse health behaviors such as high caloric intake and sedentary activity. Furthermore, only a very small proportion of the variability in BMI in this population could be attributed to socio-demographic characteristics, smoking, alcohol use, and energy intake and expenditure. There is a need to quantify more precisely the effects of energy intake and expenditure on obesity levels in populations, to determine the temporal relationship between these life-style factors and the risk of developing obesity, and to identify and clarify the roles of genetic and other environmental factors (both behavioral and psychosocial) as risk factors, especially in low socioeconomic and minority subgroups. Such research must be conducted before effective public health interventions can be designed that target the special situation of low-income individuals, particularly women. □

Acknowledgments

This research is supported by Grant HL33211 from the National Heart, Lung, and Blood Institute.

Portions of this paper were presented at the annual meeting of the Society for Epidemiologic Research, June, 1991, Buffalo, NY.

The authors thank Steven R. Browning, Joanne M. Garrett, Robert C. Kleckner, and Brian L. Rice for their valuable contributions to this work.

References

1. Burton BT, Foster WR, Hirsch J, Van Itallie TB. Health implications of obesity: an NIH consensus development conference. *Int J Obes.* 1985;9:155–159.
2. National Institutes of Health, Consensus Development Panel on the Health Implications of Obesity. Health implications of obesity. *Ann Intern Med.* 1985;103:1073–1077.
3. National Center for Health Statistics, Najjar MF, Rowland M. *Anthropometric Reference Data and Prevalence of Overweight, United States, 1976–80.* Washington, DC: US Government Printing Office; 1987. DHHS publication no. (PHS) 87-1688. (Vital and health statistics; series 11; no. 238).
4. Kumanyika S. Obesity in Black women. *Epidemiol Rev.* 1987;9:31–50.
5. Stunkard AJ. From explanation to action in psychosomatic medicine: the case of obesity. *Psychosom Med.* 1975;37:195–236.
6. Garn SM. Family-line and socioeconomic factors in fatness and obesity. *Nutr Rev.* 1986;44:381–386.
7. Flegal KM, Harlan WR, Landis JR. Secular trends in body mass index and skinfold thickness with socioeconomic factors in young adult women. *Am J Clin Nutr.* 1988;48:535–543.
8. Flegal KM, Harlan WR, Landis JR. Secular trends in body mass index and skinfold thickness with socioeconomic factors in young adult men. *Am J Clin Nutr.* 1988;48:544–551.
9. Burke GL, Jacobs DR, Sprafka JM, Savage PJ, Sidney S, Wagenknecht LE. Obesity and overweight in young adults: the CARDIA study. *Prev Med.* 1990;19:476–488.
10. Van Itallie TB. Health implications of overweight and obesity in the United States. *Ann Intern Med.* 1985;103:983–988.
11. Freedman DS, Strogatz DS, Eaker E, Joesoff MR, DeStefano F. Differences between Black and White men in correlates of high density lipoprotein cholesterol. *Am J Epidemiol.* 1990;132:656–669.
12. Strogatz DS, James SA, Haines PS, et al. Alcohol consumption and blood pressure in Black adults: the Pitt County Study. *Am J Epidemiol.* 1991;133:442–450.
13. James SA, Keenan NL, Strogatz DS, Browning SR, Garrett JM. Socioeconomic status, John Henryism, and blood pressure in Black adults: the Pitt County Study. *Am J Epidemiol.* In press.
14. Gerber AM, James SA, Ammerman AS, et al. Socioeconomic status and electrolyte intake in Black adults: the Pitt County Study. *Am J Public Health.* 1991;81:1608–1612.
15. US Department of Health and Human Services. *The Health Consequences of Smoking: Nicotine Addiction: A Report of the Surgeon General.* Washington, DC: US Government Printing Office; 1988. DHHS publication no. (CDC) 88-8406.
16. Hellerstedt WL, Jeffrey RW, Murray DM. The association between alcohol intake and adiposity in the general population. *Am J Epidemiol.* 1990;132:594–611.
17. Jequier E. Energy utilization in human obesity. *Ann NY Acad Sci.* 1987;499:73–83.
18. Metropolitan Life Insurance Co. 1983 Metropolitan height and weight tables. *Stat Bull Metropolitan Life Insur Co.* 1983;64(1):2–9.
19. Metropolitan Life Insurance Co. Measurement of overweight. *Stat Bull Metropolitan Life Insur Co.* 1984;65(1):20–23.
20. Ainsworth BE, Keenan NL, Strogatz DS, Garrett JM, James SA. Physical activity and hypertension in Black adults: the Pitt County Study. *Am J Public Health.* 1991;81:1477–1479.
21. Block G, Hartman AM, Dresser CM, Carroll MD, Gannon J, Gardner L. A data based approach to diet questionnaire design and testing. *Am J Epidemiol.* 1986;24:452–469.
22. Haskell WL, Taylor HA, Wood PD, et al. Strenuous physical activity, treadmill exercise test performance, and plasma high-density lipoprotein cholesterol. *Circulation.* 1980;62(suppl IV):53–61.
23. Holt MM; revised by Shah BV. *SUREGR: Standard Errors of Regression Coefficients from Sample Survey Data.* Research Triangle Park, NC: Research Triangle Institute; 1982.
24. Harlan WR, Landis JR, Flegal KM, Davis CS, Miller ME. Secular trends in body mass in the United States, 1960–1980. *Am J Epidemiol.* 1988;128:1065–1074.
25. Gillum RF. Overweight and obesity in Black women: a review of published data from the National Center for Health Statistics. *J Natl Med Assoc.* 1987;79:865–871.
26. Fulwood R, Abraham S, Johnson C. *Height and Weight of Adults Ages 18–74 by Socioeconomic and Geographic Variables, United States.* Washington, DC: US Government Printing Office; 1981. DHHS publication no. (PHS) 81-1674. (Vital and health statistics; series 11; no. 224).
27. Tyroler HA. Twenty-year mortality in Black residents of Evans County, Georgia. *J Clin Hypertens.* 1987;3:9s–15s.
28. National High Blood Pressure Education Program. The 1988 report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure. *Arch Intern Med.* 1988;148:1023–1038.
29. Heliovaara M, Aromaa A. Parity and obesity. *J Epidemiol Community Health.* 1981;35:197–199.
30. Noppa H, Bengtsson C. Obesity in relation to socioeconomic status: a population

study of women in Goteborg, Sweden. *J Epidemiol Community Health.* 1980;34: 139–142.
31. Forster JL, Bloom E, Sorensen G, Jeffrey RW, Prineas RJ. Reproductive history and body mass index in Black and White women. *Prev Med.* 1986;15:685–691.
32. Jeffrey RW, Forster JL, Folsom AR, Luepker RV, Jacobs DR, Blackburn H. The relationship between social status and body mass index in the Minnesota Heart Health Program. *Int J Obes.* 1989;13:59–67.
33. Khan HS, Williamson DF, Stevens JA. Race and weight change in US women: the roles of socioeconomic and marital status. *Am J Public Health.* 1991;81:319–323.
34. Albanes D, Jones Y, Micozzi MS, Mattson ME. Associations between smoking and body weight in the US population: analysis of NHANES II. *Am J Public Health.* 1987;77:439–444.
35. Shah M, Jeffrey RW, Hannan PJ, Onstad L. Relationship between socio-demographic and behaviour variables, and body mass index in a population with high-normal blood pressure: Hypertension Prevention Trial. *Eur J Clin Nutr.* 1989;43:583–596.
36. Gruchow HW, Sobocinski KA, Barboriak JJ, Scheller JG. Alcohol consumption, nutrient intake and relative body weight among US adults. *Am J Clin Nutr.* 1985; 42:289–295.
37. Williamson DF, Forman MR, Binkin NJ, Gentry EM, Remington PL, Trowbridge FL. Alcohol and body weight in United States adults. *Am J Public Health.* 1987; 77:1324–1330.
38. Frezza M, DiPadova C, Pozzato G, Terpin M, Baraona E, Lieber CS. High blood alcohol levels in women: the role of decreased gastric alcohol dehydrogenase activity and first-pass metabolism. *N Engl J Med.* 1990;322:95–99.
39. Baecke JAH, Burema J, Frijters JER, Hautvast JGAJ, van der Wiel-Wetzels WAM. Obesity in young Dutch adults: II. Daily life-style and body mass index. *Int J Obes.* 1983;7:13–24.
40. Kromhout D. Energy and macronutrient intake in lean and obese middle-aged men (the Zutphen Study). *Am J Clin Nutr.* 1983;37:295–299.
41. Ravussin E, Lillioja S, Knowler WC, et al. Reduced rate of energy expenditure as a risk factor for body-weight gain. *N Engl J Med.* 1988;318:467–472.

Postgraduate Diabetes Symposium to Be Held in Fargo, North Dakota

"Diabetes, A Rural State's Perspective—Turning Obstacles into Opportunities" is the title of the 17th Annual Symposium for physicians and other health care professionals to be held Friday, November 6, 1992, at the Holiday Inn in Fargo, ND. This year's program is a collaborative effort involving many groups, all of whom are interested in providing postgraduate education. The sponsors include: the American Diabetes Association, North Dakota Affiliate, Inc.; the North Dakota Diabetes Control Program, State Department of Health; the Centers for Disease Control; the Indian Health Service; the University of North Dakota School of Medicine; and the Fargo Veterans Administration Center.

North Dakota ranks first in the nation in the incidence of diabetes per capita. In fact, the North Dakota diabetes death toll has reached its highest level in four decades, and the state is leading the nation in deaths per capita. Because of the state's rural setting, access and availability of health care presents serious issues for health professionals caring for persons with diabetes. Therefore, the symposium will focus on these areas and also on turning obstacles into opportunities.

The program features outstanding authorities on such subjects as type II diabetes management regimes, pregnancy and diabetes, access and availability of care, complication prevention strategy in rural settings, current and future research, treatment of type I diabetes on an outpatient basis, diabetic diet and rural access to nutrition, diabetes educators' role in rural diabetes, and North Dakota outreach programs.

The registration fees to attend the symposium are $25 for members and $35 for nonmembers. Students, residents, and interns may attend for $15. The fee includes lunch, a syllabus of the program, and other educational materials. The annual meeting and awards banquet of the American Diabetes Association will be held on Thursday evening, November 5, at the same location and is open to everyone.

For more information, contact Mary Ann Keller, RN, MS, CDE, Executive Director, at the American Diabetes Association, North Dakota Affiliate, Inc., 101 North Third St, Suite 400, Grand Forks, ND 58203; tel. (701) 746-4427 or toll free 1-800-666-6709.

Prevention of Cigarette Smoking through Mass Media Intervention and School Programs

Brian S. Flynn, ScD, John K. Worden, PhD, Roger H. Secker-Walker, MB, BChir, Gary J. Badger, MS, Berta M. Geller, EdD, and Michael C. Costanza, PhD

ABSTRACT

Objectives. In this study we tested the ability of mass media interventions to enhance the efficacy of school cigarette smoking prevention programs.

Methods. For 4 years, students in one pair of communities received media interventions and school programs that had common educational objectives. Students in a matched pair of communities received only the school programs. The combined cohort of 5458 students was surveyed at baseline in grades 4, 5, and 6 and was followed up annually for 4 years.

Results. Significant reductions in reported smoking, along with consistent effects on targeted mediating variables, were observed for the media-and-school group. For cigarettes per week the reduction was 41% (2.6 vs 4.4); for smoking cigarettes yesterday the reduction was 34% (8.6% vs 13.1%); and for smoking in the past week the reduction was 35% (12.8% vs 19.8%). No effects were observed for substance use behaviors not targeted by the interventions.

Conclusions. These results provide evidence that mass media interventions are effective in preventing cigarette smoking when they are carefully targeted at high-risk youths and share educational objectives with school programs. (*Am J Public Health.* 1992;82;827–834)

Introduction

Cigarette smoking prevention programs based in middle schools have consistently shown initial reductions in the proportion of students who begin regular smoking.[1] However, when students are followed up through grade 12 these effects have been lost.[2,3] Thus, these programs may have only delayed the onset of cigarette smoking. Strategies that could be used to broaden and extend the impact of these prevention programs include enhancing their effect by combining them with educational channels outside of the school setting.[4–6] This study was designed to test the additional effect that could be achieved by adding a mass media intervention to a school smoking prevention program.

Methods

Our hypothesis was that regular cigarette smoking would increase less among adolescents receiving a mass media intervention and a school smoking prevention program that had common educational objectives than among those receiving a school program only. Each treatment group included students from one community in the Northeastern United States and one community in Montana. Five annual in-school surveys were conducted in these four communities. The surveys were linked longitudinally.

Educational Interventions

On the basis of theory[7–12] and research on smoking initiation, we developed four common objectives for the educational interventions. As a result of these interventions, young people will have:

1. a more positive view of nonsmoking;
2. a more negative view of smoking;
3. improved skills for refusing cigarettes;
4. a better understanding that most people their age do not smoke cigarettes.

Our conceptual approach linked these objectives to mediating and outcome variables (Figure 1).[7,13] Additional common objectives concerning cessation skills and awareness of cigarette marketing to young people were added during the fourth year as a result of program development research.[14]

The mass media intervention consisted of radio and television spots designed to accomplish the objectives listed above. Six different producers developed 36 television and 17 radio spots, which were broadcast as paid ads over local media. About 15 different television spots and 8 different radio spots were broadcast during each of the 4 years.

Brian S. Flynn, John K. Worden, and Roger H. Secker-Walker are with the Office of Health Promotion Research and the Vermont Cancer Center; Gary J. Badger is with the Medical Biostatistics Unit; Berta M. Geller is with the Office of Health Promotion Research; and Michael C. Costanza is with the Medical Biostatistics Unit and the Vermont Cancer Center—all at the College of Medicine, University of Vermont, Burlington, Vt.

Requests for reprints should be sent to Brian S. Flynn, ScD, Office of Health Promotion Research, 235 Rowell, College of Medicine, University of Vermont, Burlington, VT 05405.

This paper was submitted to the Journal June 13, 1991, and accepted with revisions November 12, 1991.

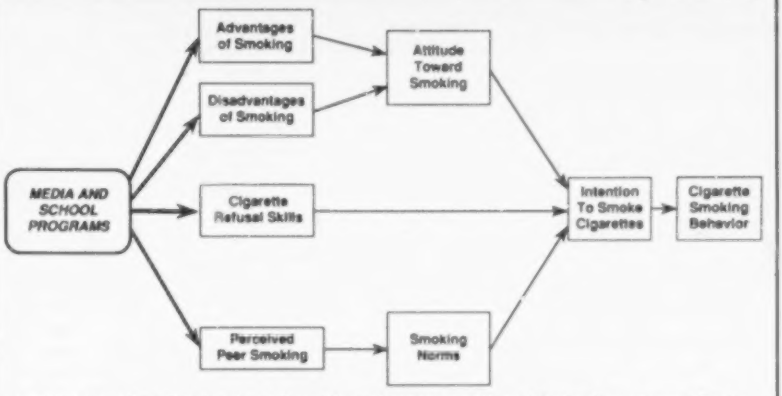


FIGURE 1—Hypothesized relationships of mass media and school program effects, mediating variables, and behavioral outcomes

The exposure plan for the first intervention year[14] was modified to match changes in media use as the target cohort matured. From the first to the fourth year, annual paid broadcast television exposures in each market were reduced from 284 to about 98; annual MTV and other cable television channel exposures were reduced from 450 to about 250; and annual paid radio exposures increased from 248 to about 450. In addition, we obtained approximately 50% further exposure from public service matching in all media. Placement of both paid and public service exposures in specific broadcast programs or viewing periods was guided by survey data.

To provide a uniform school smoking prevention program, we developed a program for grades 5 to 10 that included grade-specific curricular materials, annual teacher training, and monitoring of program implementation. The school program was based on the common educational objectives. Specific units emphasized decision making, skills to resist peer pressures and advertising pressures, social support for nonsmoking, and information about smoking and health. The curriculum was delivered by usual classroom teachers who were trained by project staff in four annual day-long teacher training workshops in each study community. For grades 5 to 8, the curriculum required four class periods per year; for grades 9 and 10, three periods were required.

The school program was linked with the media intervention only at the level of educational objectives. There was no direct programmatic relationship between the two because of the strategic decision to present the interventions as multiple, independent sources of information. The first school curriculum units were completed during the fall of 1985. The first year's television and radio campaign ran from January through May 1986; the campaign then continued at decreased intensity through the summer on radio only, with a short surge on both television and radio at the beginning of school in the fall. This pattern was repeated for 3 additional years. (Information about a videotape of selected media messages and about the school curriculum can be obtained from the authors.)

Study Sample

Two pairs of standard metropolitan statistical areas (SMSAs) were selected as best matches from regional pools. Selection criteria included being an independent media market; having a population between 50 000 and 400 000; and having matched demographic characteristics including education, income, and ethnicity. The regions were originally selected for demographic similarity to a Vermont SMSA.

One pair of SMSAs was located in the Northeastern United States (Vermont and south-central New York State) and the other was located in Montana (Table 1). Study samples were matched more closely through selection of specific school districts and feeder school units within these four SMSAs on the basis of demographic data from census tracts for the localities serving as catchment areas for the schools. Selection at this level focused on higher risk populations as indicated by adult educational attainment and household income. Initial consent to participate was obtained at the school-district level; no school refused. Within one community pair, the media-and-school community was so designated because of a need in the other community to avoid contaminating other health promotion studies within reach of the media. In the other community pair, the media-and-school community was so designated because its media served fewer outlying communities and broadcast costs were lower. A cohort of 5458 students from these four SMSAs was established in grades 4, 5, and 6 in the spring of 1985.

Measurement

Measures for each construct depicted in Figure 1 were included in each survey with the exception of refusal skills, which are difficult to measure effectively and efficiently in a questionnaire format. All measures were coded such that higher scores reflected increased levels of smoking behavior or risk.

The Smoking Behavior Index (SBI) consisted of three items that were weighted to the equivalent number of cigarettes smoked per week, then averaged.[15] The internal consistency reliability coefficient (Cronbach's alpha[16]) for this scale averaged .87 over five surveys. These self-reports were collected at the same time that a saliva sample was obtained. The students were told that the samples were taken to verify their self-reports; this procedure has been shown to enhance the validity of self-reports.[17]

"Intention to smoke" was measured by a single item. "Attitude toward smoking" was measured by a three-item scale[18] (α = .84). "Advantages of smoking" was measured by a seven-item scale[19] (α = .75). "Disadvantages of smoking" was measured by four items combined as an index. "Smoking norm" was measured by a three-item scale[20] (α = .72). "Perceived smoking by peers" was measured by a three-item scale (α = .70).

Measures of risk factors for smoking that were not targets of the intervention program were included to provide tests of the specificity of intervention effects.

Analysis

The results for the outcome measures are presented here in terms of means (proportions for dichotomous measures), with their 95% confidence intervals, for the subset of participants who took part in all five school surveys. The unit of analysis for these presentations is the individual, according to the original design of the experiment. A supplemental analysis is presented in which the community is the unit, although the original design was not intended to support this analytic strategy.

TABLE 1—Comparison of Selected Characteristics of Study Communities

	High School Graduates, %	College Graduates, %	Household Income, $	Other Language At Home, %	Non-White, %	Native To State, %	Total Population
Northeastern							
MS	71.7	16.9	17 021	7.7	2.2	74.8	173 588
SO	70.2	16.3	14 315	8.3	1.4	67.9	50 859
Montana							
MS	75.2	17.4	16 050	5.2	4.2	55.6	80 696
SO	76.6	19.9	17 460	5.2	5.8	55.2	108 035
Media-and-school	73.2	17.1	16 616	6.7	2.8	66.8	254 284
School-only	73.6	18.2	15 985	6.6	4.4	61.2	158 894

Note. MS = media-and-school; SO = school-only.
Source: 1980 US Census

Results

Less than one percent of the parents requested that their children not take part, and nearly all absentees were surveyed through later visits to schools. In this report we focus on the students who participated in all four additional surveys. Forty-seven percent (2540) of the cohort was present for all surveys (50% in the school-and-media group and 43% in the school-only group). Of those not included in the five-survey study sample, 82% moved out of the community or transferred to private schools, 15% missed one or more intermediate surveys, and the remainder were refusals, nearly all from one district that changed its consent procedures midway through the study.

Comparison with Other Baseline Participants

To assess differences between those who participated in all five surveys and the 2918 who did not, we compared baseline characteristics. Students in the five-survey sample were slightly younger (10.7 vs 10.9 years; $P = .001$) and included more girls (49.6% vs 46.4%; $P = .02$). They reported less smoking among family members, with fewer reporting any older brothers who smoked (13.8% vs 21.0%; $P < .001$), any older sisters who smoked (14.7% vs 24.9%; $P < .001$), and smoking by either parent (53.2% vs 65.3%; $P < .001$). These students also reported less smoking behavior than did those in the other group, as measured by the SBI (0.06 vs 0.53 cigarettes per week; $P < .001$), smoking yesterday (0.6% vs 3.0%; $P < .001$), and smoking in the past week (1.4% vs 5.6%; $P < .001$).

Baseline Equivalence of Treatment Groups

We compared the characteristics of those completing five surveys in the school-and-media group and in the school-only group. The school-and-media group was slightly younger (10.7 years vs 10.8 years; $P = .002$) and included more females (51.8% vs 47.0%; $P = .02$). The school-and-media group was somewhat more likely to report having an older brother who smoked (15.8% vs 11.3%; $P = .06$), but there were no other differences in family smoking. Measures of behavioral and psychosocial impact were examined to assess baseline equivalence. None of these characteristics differed significantly for the two groups, except that the school-only group reported perceiving more peer smoking (see Year 1 data in Table 2).

Impact on Cigarette Smoking Behaviors

Figure 2 depicts the SBI with 95% confidence intervals for the two study groups over 5 years. These data show a consistent trend toward less smoking by the school-and-media group, with significant differences appearing in the final 2 years. In the fifth year the school-and-media group reported smoking an average of 2.6 cigarettes per week, compared with 4.4 cigarettes per week for the school-only group, a relative difference of 41%. SBI measures for each of the four study communities are shown in Figure 3; the smoking patterns were similar within each pair of study communities. Examination of responses to individual items in this scale, indicating whether the respondent smoked any cigarettes yesterday or in the past week, showed similar patterns (Table 3). In the fifth year, smoking yesterday was reported by 8.6% of the school-and-media group and 13.1% of the school-only group, a relative difference of 34%; the percentages reporting smoking in the past week were 12.8% and 19.8%, respectively, a relative difference of 35%.

As an alternate approach to assessment of smoking behavior changes, a repeated measures analysis of variance was done using the community as the unit of analysis and blocking on geographic area (east, west). This analysis showed a significant difference between treatment groups in the trends observed over time for the SBI scores ($P < .01$).

Impact on Targeted Mediating Variables

The impact of the interventions on hypothesized mediating variables (see Figure 1) is shown in Figure 4 and Table 2. The pattern was similar for each measure. The groups were equivalent at baseline, then showed a significant difference in the second year, which persisted through the fifth year. Results from each community for three key mediating variables (Table 4) indicated similar patterns of effect within each community pair.

Impact on Variables that Were Not Targeted

Use of two substances not targeted by the interventions, smokeless tobacco and alcohol, was measured in each survey (Table 5). In smokeless tobacco behavior, reported as any use in the past week, the two groups did not differ significantly except in the fourth year, when the school-only group was more likely to report use. The groups showed no difference in alcohol use (having had beer more than once)

TABLE 2—Mean Scores for Mediating Variables Targeted by Interventions Over 5 Years, with 95% Confidence Limits

	Year	Media-And-School Group			School-Only Group		
		Mean Score	Lower Limit	Upper Limit	Mean Score	Lower Limit	Upper Limit
Smoking norms scale (range: 3–11)	1	3.64	3.57	3.70	3.77	3.69	3.85
	2	3.87	3.80	3.95	4.14[a]	4.05	4.24
	3	4.23	4.14	4.32	4.60[a]	4.50	4.73
	4	4.68	4.58	4.78	5.27[a]	5.15	5.40
	5	4.94	4.82	5.05	5.56[a]	5.43	5.68
Perceived peer smoking scale (range: 3–15)	1	5.08	4.97	5.18	5.32[a]	5.19	5.45
	2	6.06	5.94	6.17	6.42[a]	6.28	6.57
	3	6.64	6.52	6.76	7.45[a]	7.30	7.60
	4	7.44	7.32	7.56	8.71[a]	8.57	8.85
	5	8.01	7.89	8.13	9.02[a]	8.90	9.15
Attitude toward smoking scale (range: 0–3)	1	0.07	0.05	0.09	0.09	0.06	0.11
	2	0.07	0.05	0.10	0.15[a]	0.12	0.18
	3	0.10	0.08	0.12	0.22[a]	0.18	0.26
	4	0.17	0.14	0.20	0.35[a]	0.30	0.40
	5	0.25	0.21	0.29	0.37[a]	0.31	0.42
Advantages of smoking scale (range: 0–7)	1	0.44	0.39	0.49	0.42	0.36	0.47
	2	0.39	0.35	0.44	0.48	0.42	0.54
	3	0.38	0.33	0.43	0.64[a]	0.56	0.72
	4	0.52	0.45	0.59	0.78[a]	0.69	0.87
	5	0.58	0.52	0.65	0.81[a]	0.73	0.90
Disadvantages of smoking index (range: 0–7)	1	0.71	0.67	0.76	0.75	0.70	0.80
	2	0.60	0.55	0.64	0.75[a]	0.70	0.80
	3	0.58	0.53	0.62	0.78[a]	0.73	0.84
	4	0.66	0.61	0.70	0.86[a]	0.81	0.92
	5	0.69	0.64	0.73	0.79[a]	0.74	0.84

[a]95% confidence intervals around mean scores for two treatment groups do not overlap.

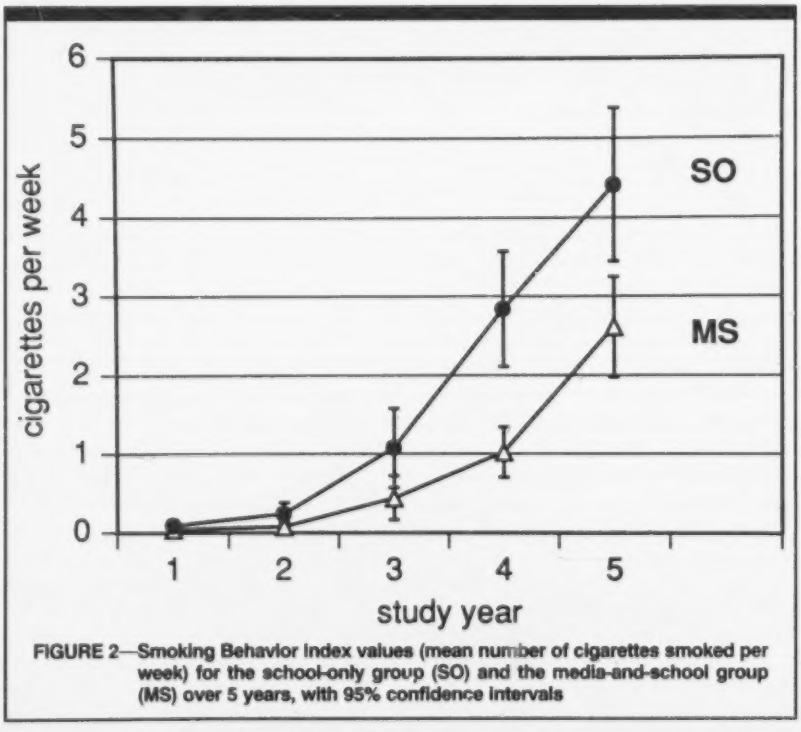


FIGURE 2—Smoking Behavior Index values (mean number of cigarettes smoked per week) for the school-only group (SO) and the media-and-school group (MS) over 5 years, with 95% confidence intervals

until higher levels of use were observed among the school-only group in the fifth year. The groups did not differ over 5 years on a psychosocial stress scale, and differed in perceived adult smoking only in year 3.

Recall of Exposure to the School Program

Items measuring recall of grade-specific elements of the school curriculum were included in the surveys for years 2 to 5. Scores were added and converted to percentages of elements recalled (Table 6). Three of the four communities produced index scores that were generally in the same range. The fourth community had somewhat lower scores.

Discussion

Consistent patterns of effect favoring the school-and-media group were found for smoking behaviors and mediators of these behaviors that were targeted by the educational interventions. Behavioral and mediating variables not targeted by the interventions generally failed to show any

trend for higher levels of risk in the school-only group. Thus, the interventions appeared to have consistent effects only on those variables it had targeted, which suggests that the effects were independent of any possible underlying general differences in study group characteristics.

The consistency of the impact of the program also can be seen clearly for smoking behavior when the two pairs of communities were viewed as four separate communities (Figure 3). The two school-and-media communities showed a clear separation from the two school-only communities in the third through the fifth surveys. Similar patterns are also seen in the data from all four communities for mediating variables (Table 4). These data also tend to support the conclusion that the effects observed were due to the interventions.

Further evidence for the internal consistency of these results can be seen by comparing the temporal ordering of the appearance of differences among the various targeted variables. In keeping with theoretical expectations, differences generally appear sooner in mediating variables (Figure 4 and Table 2) and later in smoking behavior variables (Figure 2 and Table 3).

The media-and-school group was slightly younger than the school-only group and contained a larger proportion of females. Younger age and greater female representation may offset one another as risk factors for adoption of cigarette smoking. Beginning at about grade 8, girls in this study were more likely to report that they were smokers. Separate analyses done within gender and grade-level categories (not shown here) showed patterns of effect similar to those seen in the aggregated data. Baseline comparisons of behavioral and psychosocial variables showed no differences (Table 2) except for perceived peer smoking. This difference was likely due to the slightly older age and greater male representation in the school-only group. Males were more likely to report smoking in the younger grade levels surveyed at baseline.

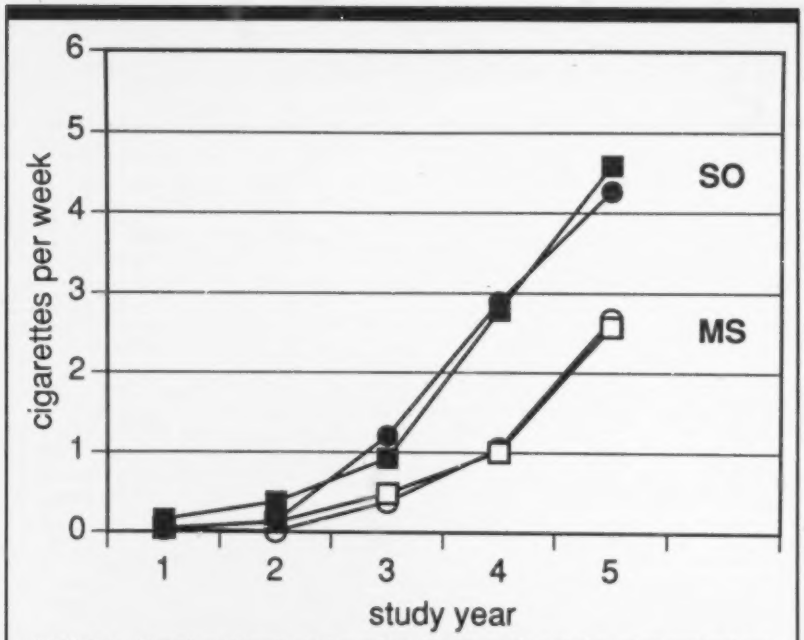


FIGURE 3—Smoking Behavior Index values (mean number of cigarettes smoked per week), by four individual study communities. The school-only communities (SO) are indicated by filled symbols and the Northeastern communities are indicated by square symbols

With respect to the school program, it should be noted that its length, starting point, and content rank it as a relatively intensive smoking prevention intervention.[1] Although differences among communities in levels of school program exposure were reported, these differences were small in relation to the differences observed in smoking behavior. Differences in reported exposure generally did not coincide with changes in smoking behavior (Figure 3 and Table 6) or mediating variables (Figure 4 and Table 4); this is especially apparent when data for the two school-only communities are compared. It is possible that these differences in level of implementation had little or no effect on smoking behavior because the lowest level of reported school program implementation exceeded a ceiling above which no further effect could have been attained.[1,4,21,22]

TABLE 3—Percentages of Students Smoking Daily or Weekly, by Intervention Group over 5 Years, with 95% Confidence Limits

		Media-And-School Group			School-Only Group		
	Year	Mean Score	Lower Limit	Upper Limit	Mean Score	Lower Limit	Upper Limit
Smoked any cigarettes yesterday	1	0.61	0.19	1.03	0.49	0.06	0.92
	2	0.76	0.29	1.24	1.27	0.58	1.96
	3	1.91	1.17	2.65	3.71	2.55	4.87
	4	5.35	4.13	6.57	9.57[a]	7.77	11.37
	5	8.56	7.05	10.08	13.09[a]	11.02	15.15
Smoked any cigarettes last week	1	1.29	0.70	1.88	1.59	0.87	2.32
	2	1.80	1.10	2.50	3.44	2.37	4.50
	3	5.01	3.86	6.15	9.25[a]	7.56	10.95
	4	9.10	7.59	10.61	14.82[a]	12.74	16.91
	5	12.61	11.04	14.57	19.80[a]	17.46	22.15

[a]95% confidence intervals around mean scores for two treatment groups do not overlap.

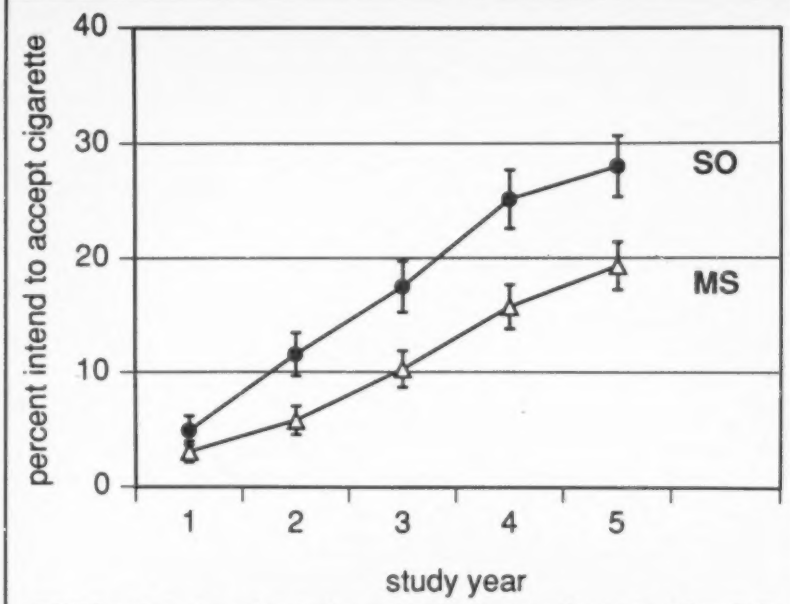


FIGURE 4—Percentage of students reporting that they intend to smoke a cigarette if offered by a friend, for the school-only group (SO) and the media-and-school group (MS) over 5 years, with 95% confidence intervals

TABLE 4—Mean Scores for Key Mediating Variables from All Four Communities over 5 Years

		Year					
	Community	1	2	3	4	5	n
Intentions (range: 0–1)	MS (NE)	0.04	0.07	0.11	0.17	0.21	831
	SO (NE)	0.04	0.08	0.13	0.20	0.25	486
	MS (M)	0.02	0.04	0.10	0.13	0.17	529
	SO (M)	0.06	0.15	0.21	0.30	0.30	621
Norms (range: 3–11)	MS (NE)	3.71	3.95	4.27	4.88	4.94	839
	SO (NE)	3.72	4.06	4.34	5.00	5.28	488
	MS (M)	3.52	3.75	4.16	4.68	4.93	520
	SO (M)	3.81	4.21	4.81	5.49	5.77	609
Attitudes (range: 0–3)	MS (NE)	0.07	0.08	0.09	0.16	0.24	825
	SO (NE)	0.09	0.13	0.15	0.23	0.29	473
	MS (M)	0.06	0.07	0.11	0.19	0.27	515
	SO (M)	0.08	0.16	0.28	0.45	0.43	606

Note. MS = media-and-school; SO = school-only; NE = Northeastern; M = Montana.

The study was designed with individual participants as the unit of analysis. However, because of the need to apply the educational programs to larger aggregates rather than to randomly selected individuals, participants were allocated to treatment conditions by communities. We have carried out the analyses according to the original design, using individuals as the primary analytic unit. An additional analysis was conducted with the community as the analytic unit, although the original design did not anticipate this approach. This supplementary analysis produced results similar to those of the main analysis, providing confirmation at the community level of the impact of the mass media intervention.

The results reported in this paper were for students who participated in all surveys. Most of the other students left the school systems after the first survey. Very few parents or students refused to participate in any survey, and nearly all absentees at the time of the main surveys were subsequently surveyed. It seems likely that those excluded from these analyses live in less stable family environments; this likelihood reduces the generalizability of these results for a part of the adolescent population that is at somewhat higher risk for adoption of smoking.[23] We are following up the entire cohort to assess the sustained impact of the program.

A second factor that limits the generalizability of the study is the lack of racial and ethnic diversity in the study population. This feature of the study is a result of a matching process that began in a relatively homogeneous area. The challenge for researchers is to adapt the strategies used here to devise equally effective campaigns for other major population groups and for more heterogeneous populations.

Interpretation of Mass Media Effect

Aside from the fact that it shared common objectives with a school program, the impact of this mass media intervention may be attributed to several features that set it apart from previous efforts using mass media in health promotion[12,14]: the intervention had sufficient duration; it used multiple channels; it was presented at times and places when young people reported their highest use of media; it used a variety of message styles; it used appeals to age- and gender-specific motives determined through surveys and focus groups; and it used messages portraying perceived social support with attractive age- and gender-relevant models providing appropriate behavioral skills, alternatives, and reinforcement.

This intervention also avoided the use of exhortations; the spokespersons were credible young models saying "I don't smoke," letting each youngster make his or her own decision. Program formats that appealed to young people were used for these messages: situation comedies, cartoons, rock videos, and testimonials. Because young people at higher risk for smoking also tend to shun authority, the messages were not "tagged" with a logo or name of a sponsoring agency; the sponsor was identified only in an unobtrusive copyright notice. Lastly, rather than presenting only a few spots in each "flight" (6- to 15-week period) in an effort to ensure awareness of specific content, we presented 12 to 18 different spots in each flight to establish a perception of a nonsmoking lifestyle; more popular spots were repeated annually.

It is important to underscore the role that the media play in the socialization of

TABLE 5—Mean Scores for Variables Not Targeted by Intervention over 5 Years, with 95% Confidence Limits

		Media-And-School Group			School-Only Group		
	Year	%	Lower Limit	Upper Limit	%	Lower Limit	Upper Limit
Used smokeless tobacco in past week, %	1	2.66	1.79	3.53	2.54	1.59	3.49
	2	3.57	2.57	4.57	4.24	3.03	5.45
	3	3.34	2.37	4.31	5.09	3.77	6.41
	4	5.16	3.96	6.36	8.29[a]	6.63	9.95
	5	8.28	6.79	9.77	9.61	7.84	11.38
Beer drinking more than once, %	1	7.36	5.97	8.75	8.58	6.91	10.25
	2	15.53	13.59	17.47	16.26	14.06	18.46
	3	24.74	22.43	27.05	27.94	25.27	30.61
	4	41.01	38.38	43.64	42.01	39.09	44.93
	5	49.48	46.81	52.15	55.80[a]	52.86	58.78
Perceived adult smoking (range: 1–5)	1	3.53	3.48	3.58	3.57	3.51	3.63
	2	3.50	3.45	3.54	3.50	3.44	3.55
	3	3.37	3.33	3.41	3.47[a]	3.42	3.52
	4	3.27	3.23	3.31	3.34	3.30	3.39
	5	3.21	3.17	3.25	3.26	3.22	3.31
Life stress scale (range: 0–4)	1	0.34	0.30	0.38	0.34	0.30	0.38
	2	0.34	0.30	0.38	0.35	0.30	0.39
	3	0.34	0.30	0.38	0.39	0.34	0.44
	4	0.39	0.35	0.43	0.41	0.36	0.45
	5	0.40	0.36	0.44	0.38	0.34	0.43

[a]95% confidence intervals around mean scores for two treatment groups do not overlap.

young people. Television has been considered to be a major agent of socialization through which children learn the definitions and expectations underlying the rules and norms of their culture.[24] Parents and teachers witness this process every day as new trends seen in the media become the new behaviors to be adopted in the real world. Because the messages used in this intervention had no unifying tag or logo, and because they appeared to come from many sources simultaneously, it is possible that the young people in our intervention communities thought that these messages were being received by young people across the nation and perceived them as redefining norms for youth smoking.

Other Mass Media and Smoking Prevention Studies

The present study is one of three research projects that were funded by the National Cancer Institute in 1984 to evaluate mass media smoking prevention approaches. The others involved school and family participation with evening newscast coverage[25] and television and radio spot messages with sweepstakes participation.[26] Bauman et al.[27] reported that the media-sweepstakes program produced changes in some mediating variables.

TABLE 6—Mean School Program Exposure Index Scores (Percentages) from All Four Communities

	Year				
Community	2	3	4	5	n
MS (NE)	80.4	84.4	69.2	67.0	811
SO (NE)	76.5	80.1	69.3	58.9	480
MS (M)	79.2	81.6	65.0	66.2	516
SO (M)	69.3	76.9	49.4	49.1	610

Note. MS = media-and-school; SO = school-only; M = Montana; NE = Northeastern.

A media and school-based campaign conducted in Australia, in which an uncontrolled evaluation design was used, resulted in reductions in adolescent smoking.[28] Pentz et al. reported the impact on adolescent smoking of a multicomponent prevention program,[29] but mass media effects were not assessed independently. These two apparently successful programs combined mass media with school programs and other channels of influence for young people. They provide additional evidence that carefully targeted mass media interventions can have a favorable impact on adoption of cigarette smoking when the media interventions are combined with school programs or other powerful channels of interpersonal influence.

Larger Scale Replication

The design and implementation process used for this program could be replicated on a larger scale at a reasonable level of investment. In this study the mass media intervention was implemented in two widely separated locations, using similar techniques with similar effectiveness. The media tastes of young people in both communities were almost identical. To mount such a campaign on a larger scale would require agreement between media and school program planners on common objectives; access to small student samples for assistance with media intervention development; resources to produce media spots designed to appeal to specific target groups; and money to buy enough

broadcast time in selected markets to ensure high levels of exposure. The average cost of producing a finished television message for this study, including development costs, was $10 369. Average broadcast costs in two small-to-medium markets from 1986 to 1989 were $25 000 per market annually.

Several factors work in favor of achieving consistency between school program objectives and mass media objectives in a larger scale replication. The school program used in this study was explicitly designed in 1984 to represent a good smoking prevention program for the future. Continued strengthening of school health education curricula has occurred since then,[30] with many programs based on objectives similar to those used here. These changes set the stage for more effective collaboration between school-based and mass media interventions.

Conclusion

The interventions for this study were based on three main strategies: (1) common educational objectives for concurrent mass media and school-based interventions, (2) careful targeting of interventions, and (3) maintenance of interventions at high levels of intensity and diversity throughout the critical years for smoking adoption.

Those responsible for directing media interventions should understand the objectives and content of school-based programs being conducted in the broadcast area so that they can reinforce these activities. Media planners also should use diagnostic and formative research to learn about message styles appealing to young people at different developmental stages. Messages should be presented with adequate frequency and variety to promote a lifestyle with positive nonsmoking norms.

The strength of this approach is in the promotion of a lifestyle, rather than a product or a service. There are examples from advertising in which selling a product also promotes a lifestyle, most notably in beer advertising. Taken one at a time, beer commercials sell a product; taken together, they promote an easygoing, highly social lifestyle. The media and school interventions in this study accomplished a similar goal by presenting messages that appeared to come from different sources but had the same educational objectives.

With their appeal to young people, especially those at highest risk for smoking, mass media provide a particularly effective way to deliver smoking prevention messages. The approach developed in this study apparently tapped the power of the media to influence smoking norms among this highly impressionable age group, and prevented significant proportions of them from starting to smoke cigarettes. □

Acknowledgments

This study was supported by research grants CA38395 and CA22435 awarded by the National Cancer Institute and research grant HL17292 awarded by the National Heart, Lung, and Blood Institute.

References

1. Glynn T. Essential elements of school-based smoking prevention programs. *J Sch Health.* 1989;59:181–188.
2. Murray D, Davis-Hearn M, Goldman A. Four- and five-year follow-up from four seventh grade smoking prevention strategies. *J Behav Med.* 1988;11:394–405.
3. Flay B, Koepke D, Thomson S. Six-year follow-up of the first Waterloo school smoking prevention trial. *Am J Public Health.* 1989;79:1371–1376.
4. Flynn B, Worden J, Secker-Walker R. Cigarette smoking prevention: constraints and alternatives. In: Health Education Bureau, eds. *Proceedings of the 12th World Conference on Health Education, Dublin, Ireland, 1987;* 505–507. Available from the International Union for Health Education, Health Education Bureau, Dublin, Ireland.
5. Worden J, Flynn B, Brisson S, Secker-Walker R, McAuliffe T, Jones R. An adult communication skills program to prevent adolescent smoking. *J Drug Educ.* 1987; 17:1–9.
6. Rice RE, Paisley WJ, eds. *Public Communication Campaigns.* Beverly Hills, Calif: Sage; 1981:7–11.
7. Fishbein M, Ajzen I. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research.* Reading, Mass: Addison-Wesley; 1975.
8. Jessor R, Jessor S. Problem Behavior and Psychosocial Development: A Longitudinal Study of Youth. New York, NY: Academic Press; 1977.
9. Jessor R. Critical issues in research on adolescent health promotion. In: Coates T, Petersen A, Perry C, eds: *Promoting Adolescent Health: A Dialog on Research and Practice.* New York, NY: Academic Press; 1982:447–465.
10. Bandura A. Self-efficacy: toward a unifying theory of behavior change. *Psychol Rev.* 1977;84:191–215.
11. Akers R, Krohn M, Lanza-Kaduce L, Radosevich M. Social learning and deviant behavior: a specific test of a general theory. *Am Sociological Rev.* 1979;44:636–655.
12. Flay BR, DiTecco D, Schlegel R. Mass media in health promotion: an analysis using an extended information-processing model. *Health Educ Q.* 1980;7:127–147.
13. Green LW, Kreuter MW. *Health Promotion Planning: An Educational and Environmental Approach.* Mountain View, Calif: Mayfield; 1991.
14. Worden J, Flynn B, Geller B, et al. Development of a smoking prevention mass media program using diagnostic and formative research. *Prev Med.* 1988;17:531–558.
15. Pechacek T, Murray D, Luepker R, Mittelmark M, Johnson C, Shutz J. Measurement of adolescent smoking behavior: rationale and methods. *J Behav Med.* 1984;7:123–140.
16. Nunnally J. *Psychometric Theory.* New York, NY: McGraw-Hill; 1967.
17. Murray D, Perry C. The measurement of substance use among adolescents: when is the bogus pipeline method needed? *Addict Behav.* 1987;12:225–233.
18. Chassin L, Corty E, Presson C. Predicting adolescents' intentions to smoke cigarettes. *J Health Soc Behav.* 1981;22:445–455.
19. McAlister A, Krosnick J, Milburn M. Causes of adolescent cigarette smoking: tests of a structural equation model. *Soc Psychol Q.* 1984;47:24–36.
20. Krohn M, Akers R, Radosevich M, Lanza-Kaduce L. Norm qualities and adolescent drinking and drug behavior: the effect of norm quality and reference group on using and abusing alcohol and marijuana. *J Drug Issues.* 1982;12:343–359.
21. Flay BR. Psychosocial approaches to smoking prevention: a review of findings. *Health Psychol.* 1985;4:449–488.
22. Best AJ, Thomson SJ, Santi SM, Smith EA, Brown KS. Preventing cigarette smoking among school children. *Ann Rev Public Health.* 1988;9:161–201.
23. Pirie PL, Murray DM, Luepker RV. Smoking prevalence in a cohort of adolescents, including absentees, dropouts, and transfers. *Am J Public Health.* 1988;78:176–178.
24. Comstock G, Chaffee S, Katzman N, McCombs M, Roberts D. *Television and Human Behavior.* New York, NY: Columbia University Press; 1978.
25. Flay BR, Brannon B, Johnson C. The television school and family smoking prevention and cessation project: theoretical basis and program development. *Prev Med.* 1988;17:585–607.
26. Bauman K, Padgett C, Koch G. A media-based campaign to encourage personal communication among adolescents about not smoking cigarettes: participation, selection, and consequences. *Health Educ Research.* 1989;4:35–44.
27. Bauman KE, LaPrelle J, Brown JD, Koch GG, Padgett CA. Influence of three mass media campaigns on variables related to adolescent smoking: results of a field experiment. *Am J Public Health.* 1991; 81:597–604.
28. Pierce J. Time to ban advertising and continue the Quit For Life campaigns. *Med J Aust.* 1990;152:113.
29. Pentz M, Dwyer J, MacKinnon D, et al. A multicommunity trial for primary prevention of adolescent drug abuse: effects on drug use prevalence. *JAMA.* 1989;261: 3259–3266.
30. Task Force on Education of Young Adolescents. *Turning Points: Preparing American Youth for the 21st Century.* Washington, DC: Carnegie Council on Adolescent Development; 1989.

The Second Chicago Televised Smoking Cessation Program: A 24-Month Follow-up

Richard B. Warnecke, PhD, Patricia Langenberg, PhD, Siu Chi Wong, MS, Brian R. Flay, DPhil, and Thomas D. Cook, PhD

ABSTRACT

Objectives. As smoking decreases in the population, the remaining smoking population will change, and cessation initiatives will have to incorporate strategies designed for these smokers.

Methods. To study patterns of response to a cessation intervention composed of 20 televised segments and the American Lung Association *Freedom from Smoking in 20 Days* manual, this study compared cessation rates over 24 months in a cohort of smokers who registered for a cessation program with those in a cohort selected from the smoking population at large.

Results. At post intervention, multiple point prevalence of cessation among participants, adjusted for baseline smoking, was 14% among registrants and 6% in the population; at 24 months the adjusted rates were 6% and 2%, respectively. Heavy smokers benefited more than light smokers, and there was a consistent dose-response relationship between extent of exposure to the intervention and cessation.

Conclusion. The effects of the intervention were strongest for those who read the manual and watched the programs daily. Manual use was important, and those who did not read it did not appear to benefit. Compared to the population and given full participation, heavy smokers benefited more than lighter smokers. (*Am J Public Health.* 1992;82; 835–840)

Introduction

Data obtained in the National Health Interview Survey through 1985 show that about 30% of the adult US population over the age of 20 years smokes.[1] The goal of the US Public Health Service and the National Cancer Institute is to reduce smoking among that population to 20% by the year 2000.[2,3] Other data compiled from the survey indicate that the smoking population of the future is likely to comprise hard-core smokers who either are unmotivated to quit or have failed to succeed at past efforts.[4–6] These remaining smokers have not responded to traditional smoking interventions that have been successfully offered in the past to the general population.[7–9].

Our ongoing research and that of others indicate that these continuing smokers are likely to have needs that limit their capacity to respond to such interventions. Thus, new interventions will probably have to incorporate different strategies to reach these smokers. For example, those who continue to smoke may be most likely to respond to interventions that do not call for attendance at face-to-face meetings and that require minimal reading or provide reading materials that are written at a very low literacy level.[4,10]

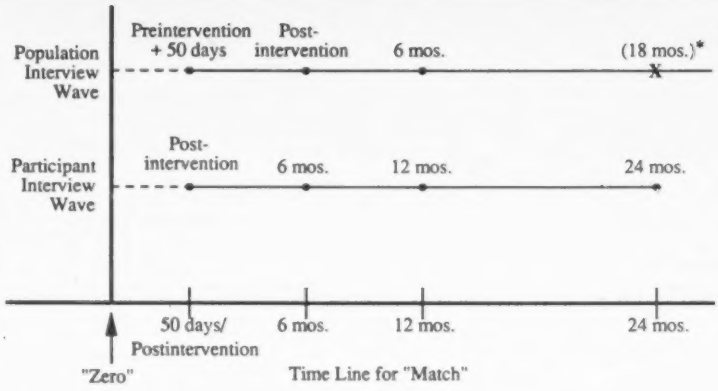
Televised smoking cessation interventions that incorporate models to whom these smokers can relate are often used because they reach many smokers. Although some community-based programs have not found televised interventions to be effective without direct, face-to-face support from community health workers,[11,12] others have found otherwise, especially when such interventions are combined with other self-help materials.[9,13] Elsewhere[10] we reported that those at risk to continue smoking were likely to view credible televised programs, such as those offered as part of a news broadcast, and could recall 6 or more of the 20 segments that were broadcast. However, analysis of participation in the intervention also indicated that those who recalled segments of the broadcasts were not the same as those who referred to the accompanying manual. Thus, if the effectiveness of intervention is contingent both on viewing and recalling televised segments and on referring to a manual, the intervention may not fully reach those likely to be the smokers in the future. As we show here, however, the combination of referring to a manual *and* viewing the programs daily predicts the greatest success with the program.

The principal purpose of this paper is to report long-term (24-month) results from the minimal self-help smoking cessation intervention that combined the American Lung Association manual, *Freedom from Smoking in 20 Days,* with a daily series of 20 televised news segments designed to supplement the daily steps described by the manual. These results compare the 24-month multiple point

Richard B. Warnecke is with the Survey Research Laboratory, Brian R. Flay is with the Prevention Research Center, and Siu Chi Wong is with the Epidemiology-Biostatistics Program at the University of Illinois at Chicago. Patricia Langenberg is with the Department of Epidemiology and Preventive Medicine at the University of Maryland at Baltimore, and Thomas D. Cook is with the Sociology Department at Northwestern University in Evanston, IL.

Requests for reprints should be sent to Dr. Richard B. Warnecke, Survey Research Laboratory, University of Illinois at Chicago, P.O. Box 6905, M/C 336, Chicago, IL 60680.

This paper was submitted to the Journal December 26, 1990, and accepted with revisions November 27, 1991.



*Status determined from 24-month interview.

FIGURE 1—Lagged analysis plan

prevalence of cessation of a cohort of program participants with that of a sample drawn from the smoking population targeted by the intervention. A second purpose is to describe how level of exposure to the intervention affected rate of sustained cessation and timing of the cessation event.

Methods

Design of the Intervention

Details of the intervention and characteristics of the participants are described in detail elsewhere.[10] Briefly, the televised portion of the intervention was broadcast over 20 days as part of the local evening news at 4 PM and 10 PM on WLS-TV, the local ABC-owned station in Chicago. During the 3-week registration period, 10- and 30-second spots, featuring a mime, "kicking the habit," and donated by WLS-TV, were broadcast telling smokers how to register. Through collaboration with the Chicago Lung Association, True Value Hardware stores, and PruCare Health Maintenance Organization, the American Lung Association manual, *Freedom from Smoking in 20 Days,* was distributed at no cost when smokers registered for the program or requested the manual. In all, 9182 smokers registered for the program, and an estimated 75 000 additional manuals were distributed through the stores to those smokers who requested them.

Study Samples

The results reported here are from a sample of 2089 participants and a sample of 2398 smokers in the general population who reported viewing the evening news at least 3 days a week and were not participants. Participants were registrants who, at the first postintervention interview, reported referring to the manual at least once and/or viewing at least one segment of the program.

In October 1986 (6 months before the intervention was broadcast), two population panels were selected through telephone screening from the population of smokers in the Chicago metropolitan statistical area. Their purpose was to provide baseline estimates of population smoking behavior prior to the intervention and to model changes in smoking behavior in that population absent any intervention. Because the two panels were selected to be equal in number, a disproportionate, stratified design was used with directory-based random-digit dialing. We screened 35 541 telephone numbers to obtain 2398 completed interviews.

The population panels were selected from all smokers 18 years of age and older who resided in the six-county area comprising the Chicago metropolitan statistical area, smoked at least seven cigarettes a week, and viewed the late afternoon or evening news at least three times per week (including weekends). The two panels were distinguished by their channel of preference for viewing the evening news as determined by their responses to the initial interview and to a second preintervention interview 3 months later. Smokers who did not watch the evening news were dropped from the panels.

Statistical Methods

To compare the smoking cessation patterns among the participant and population samples over the 24 months following the intervention, adjusting for differences in demographics and baseline smoking, and to model smoking cessation patterns absent any intervention, we used the weighted, combined population data, which is described as the population panel.

A "lagged" analysis compared changes in smoking behavior among subjects in the participant sample and population panel at 50 days and at 6, 12, and 24 months past the "zero" point, defined as the *time of first interview* for the population panel and the *date of the start of the intervention* for the participants. The matching point was arbitrarily selected at 50 days because the median time interval between the intervention zero point and the first postintervention interview was 50 days (see Figure 1). At zero point, all subjects were smoking. At each interview, the exact (or as nearly exact as possible) date of "quit" was ascertained. Because the final interviews were conducted after a 12-month interval, there was no interview at 18 months for the participants or at 24 months for the population. (The population interviews began 6 months before the intervention.) The 24- and 30-month interviews included carefully designed probes for smoking status at 18 months for participants and at 24 months for the population. These data were used to estimate with reasonable confidence the quitting status at the relevant points of comparison for the 6-month point for both samples.

The analysis focused on examining and comparing cessation patterns in the participant sample and the weighted, combined population panel, overall and within intervention exposure groups. Survival analysis (event history) methodology was used to model and plot time to cessation from the zero point. Because respondents were interviewed at different times, they were either quit at the time of interview or censored (final outcome not determined) at the date of the interview. Analyses were made at each interview wave so that each analysis answers the question: "For a respondent who was not smoking at a given interview point, when did the *last* quit without subsequent relapse occur?" The respondent had to have been quit for 3 or more days at the first postintervention interview and for 7 or more days at subsequent postintervention interviews to be considered a nonsmoker. Modeling was performed using the Cox proportional hazards method to adjust for demographic and smoking history variables on which the samples differed in preliminary cross-tabulations. Because the proportional

hazards assumptions were not met for exposure subgroups, this variable was entered as a stratifying variable. Cumulative cessation rate curves (complement of survival curves) were plotted for exposure groups.

Response Rates

In all, 74% of the participants were retained through 24 months, and 51% of the population panel members were retained through 30 months. Rather extensive attrition occurred within the study period, particularly in the population panel. Because the sample was selected using random-digit dialing, in which telephone numbers are selected randomly and include unlisted and directory-assisted numbers, respondents were often lost to follow-up when they changed residence. Although we collected information about a person who would know how to reach the respondent, often this secondary source could not be located or refused to provide the information to locate the respondent.

As shown in Table 1, males, Blacks, low-income, and less-educated respondents were overrepresented among respondents lost to follow-up compared with those who remained over the entire period. Although, as expected in large samples, many of the differences in the table are statistically significant, most of the actual differences are relatively small. Of more importance, the baseline smoking rates did not differ in those followed up compared with the rates in those lost to follow-up. The distribution of smoking behavior at the time of loss to follow-up did not change among those who remained and those who were lost (not shown in the table). Respondents who were not recontacted in a given wave but were recontacted at a later wave were questioned about their smoking status for the missing wave and were retained in the continuous cohort when their smoking behavior for the missing wave could be ascertained.

TABLE 1—Demographic Characteristics and Baseline Smoking Behavior of Respondents Lost to Follow-up vs Those Followed over the Entire 24-Month Period, by Percentage

Variables	Participants		Population Panel	
	Lost to Follow-up (n = 539)	Followed over 24-Month Period (n = 1550)	Lost to Follow-up (n = 1181)	Followed over 24-Month Period (n = 1217)
Gender				
Male	33	27*	47	44
Female	67	73	53	56
Race				
Non-Black	66	78*	82	87*
Black	34	22	18	13
Income				
<$13K	29	15*	15	11*
$13K–$39K	44	49	52	53
$40K+	18	32	25	32
Unknown	9	4	8	4
Education				
<High school	15	9*	16	10*
High school	38	37	36	35
College	43	48	42	48
Graduate school	4	6	6	8
Health status				
Very poor	3	2*	2	1
Poor	13	9	6	4
Good	59	60	56	56
Very good	25	29	36	39
Average no. of cigarettes smoked daily				
0–19	26	21	44	39
20–29	35	35	35	38
30+	39	44	21	23

*$P < .01$ by χ^2.

Results

First, the intervention attracted a disproportionate number of heavy smokers (30 or more cigarettes per day) relative to the population (see Table 1). Female and Black smokers were also overrepresented among participants compared with the population.[10]

We examined point prevalence of cessation following the intervention and multiple point prevalence of cessation at subsequent observation points. Multiple point prevalence is defined here as the cessation rate at any subsequent wave of data collection among those participants and members of the population panel who reported being abstinent for at least 3 days at 50 days post zero point and who reported being abstinent for at least 7 days at each intermediate contact. It is our estimate of continuous cessation; however, as in most studies where it is used, multiple point prevalence is biased because it includes as abstinent any persons who may have slipped between measurement points.

The overall point prevalence of cessation among participants at 50 days post zero, adjusted for baseline smoking level, was 16%; the comparable point prevalence rate in the population panel was 4%. At 6 months, the adjusted multiple point prevalence of cessation was 9% for the participants and 3% for the population; at 12 months, the rates were 6% and 3%, respectively; and at 24 months, they were 6% and 2%, respectively. Relapse during the 6 months following the intervention was greater for the participants than for the population; nevertheless, the 24-month multiple point prevalence rate among the participants was three times that observed in the population.

Table 2 presents these data broken down by average number of cigarettes smoked daily at baseline. Several things are worth noting in the table. First, the point prevalence of cessation at post intervention for each level of baseline smoking is higher among participants than among the population panel, consistent with an intervention effect. Second, light smokers at baseline were more likely than heavy smokers to be abstinent at 50 days post zero. However, although the heaviest smokers among the participants quit at two thirds the rate of the light smokers, the rate of cessation among the heaviest smokers in the population was only 25% of the rate among light smokers. Third, the rate of relapse at 24 months postintervention among light smokers was 70% among the participants and 62% among the population; among the heaviest smokers, it was 60% among the participants and zero among the population. These comparisons suggest that the heaviest smokers in both groups were somewhat better able to sustain their decision to quit than the light

TABLE 2—Point Prevalence of Cessation of 50 Days/Postintervention (PI) and Multiple Point Prevalence of Cessation over 24 Months, Controlling for Number of Cigarettes Smoked at Baseline

	Average Number of Cigarettes Smoked Daily					
	0–19		20–29		30+	
	Participants, % (n = 321)	Population Panel[a] (n = 471)	Participants, % (n = 548)	Population Panel[a] (n = 466)	Participants, % (n = 681)	Population Panel[a] (n = 279)
50 days/PI	23	8*	14	4*	15	2*
PI–6 months	11	8	7	3*	9	2*
PI–12 months	8	6	5	2	7	2*
PI–24 months	7	3*	4	1*	6	2*

*$P < .05$ by χ^2.
[a]Weighted.

smokers but that the benefits of the program were most obvious in the high initial quit rates among heavy smokers who participated. That is, among light smokers the ratio of initial point prevalence of cessation of participants to that of the population was about 3 to 1 in favor of participants; the comparable ratio among heavy smokers was 7.5 to 1, also in favor of participants.

Figure 2 presents the cumulative cessation rate curves at 50 days post zero and at 6, 12, and 24 months following the intervention, adjusted for gender, race, education, smoking rates at baseline, and self-reported health status. Separate curves were plotted for various levels of participation in the intervention. Participation (n = 1550) was defined at three levels for the plots presented here: *high participation* (10%) accounts for those who watched the televised segments daily and referred to the manual daily; *moderate participation* (37%) accounts for those who watched the televised segments at least once a week and referred to the manual at least two to six times a week; and *low participation* (53%) accounts for those who participated to any lesser extent.

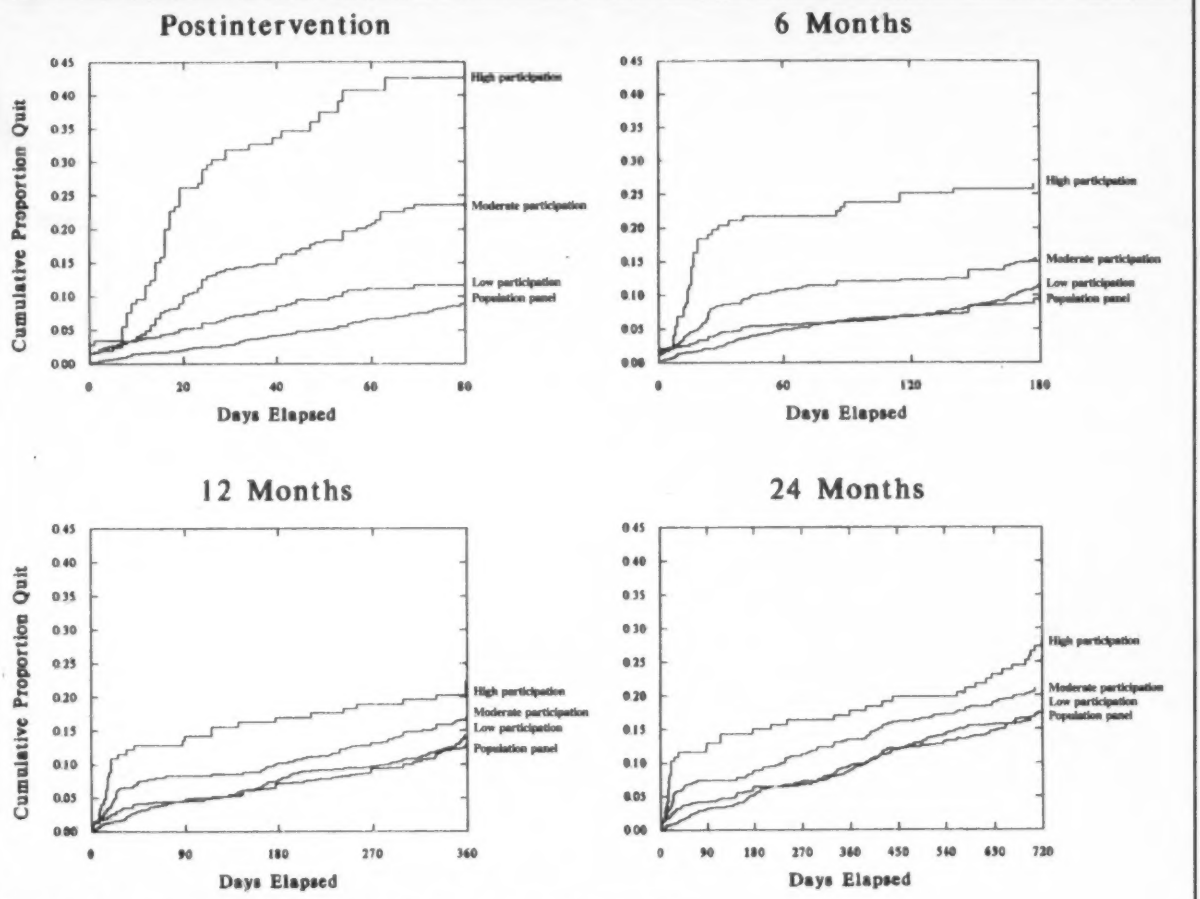
In Figure 2, at each interview wave, even at 24 months, the cumulative proportion quit among high participants was considerably greater than that among moderate participants, low participants, and the population panel. Consistent with the data in Table 2, the greatest contrast is between high participants and the population panel at every point. Also consistent with Table 2, the initial increases in slope for the high and moderate participants indicate that, among those who were abstinent at 24 months, most of the cessation occurred as part of the intervention and clustered around day 18, the quit day specified by the program. The gradual increase in the slope of the low participants is similar both to that of the population and to what would be expected with no intervention.

In Table 3, we present results of the Cox proportional hazards method, predicting cessation across waves of intervention as a function of demographic characteristics, baseline smoking behavior, and overall health status. The variables above the dotted line, which are the variables associated with cessation on which the participants and the population panel differed, are the covariates that were controlled in the Cox proportional hazards analysis. Race and education are sufficiently correlated so that neither appears to have an independent, significant association with quitting. However, having a college education *is* significantly and positively associated with cessation at postintervention compared with having less than a high school education. The association appears to reverse at 12 months. Race is significant only at 24 months, although non-Blacks were more likely than Blacks to be consistently abstinent across all waves. Males were more prone than females to have quit at all points, although the relationship is not statistically significant at 24 months. Health is also an important factor. Those in good health were more likely to be abstinent at each observation; however, the relationship is not significant at the 12-month observation. Only baseline smoking level is consistently associated with cessation through 24 months: lighter smokers were more likely to quit than heavier smokers.

Below the dotted line in Table 3 are the risk ratios associated with participation level. Because the proportional hazards assumptions were not met for the participation variable, the confidence interval estimates should be viewed with caution as they tend to average across time. The inferences support the findings from the cumulative cessation plots that those in the high and moderate participation groups were more likely to quit and remain quit than were those in the population panel, although the effects are not always significant.

Discussion

Participants who followed the program with the greatest regularity were more likely to quit smoking and to sustain their abstinence over 24 months than were those who participated less intensively or individuals in the population who quit without the intervention. Moreover, as seen from the curves in Figure 2, there was a dose-response relationship between level of participation and maintenance of abstinence over 24 months, with larger risk ratios for higher participation categories. Furthermore, the cumulative cessation rate curves show clearly that most of those who were abstinent at 24 months quit initially during the intervention period. Finally, the long-term effects of participation in the intervention were especially evident for heavy smokers, given that those who participated were more likely to quit than were those in the population panel, who did not have access to the intervention (see Table 2). Moreover, among the participants, although light smokers were more likely to quit postintervention, heavy smokers who quit were somewhat more likely to be abstinent at 24 months. Thus,



High Participation = Watched televised segments daily and referred to manual daily
Moderate Participation = Watched televised segments at least once a week and referred to manual at least two to six times a week
Low Participation = Participated to any lesser extent

FIGURE 2—Cumulative cessation rates

all the data are consistent with an intervention effect and suggest particular benefit for heavy smokers.

The target population in this study was defined as those smokers who watched the evening news on the intervention television channel. The demographic characteristics of the viewing audience indicate that the news audience for this channel contained a disproportionate number of those who are likely to be the smokers in the year 2000. These include heavy smokers, low-socioeconomic females, Blacks, and blue-collar workers.[4-6] Therefore, the fact that the participants drawn from this audience did so much better than the quitters from this audience who attempted to quit without the intervention is also quite encouraging, as are the positive effects of the program on heavy smokers who participated. This latter finding is particularly noteworthy, given that heavy smokers were represented disproportionately relative to their distribution in the population.

Based on the risk ratios in Table 3, however, light smokers at each wave of observation were most likely to be abstinent. Although the relationship is not consistently significant, males appear to have had more success with stopping than females, as did smokers in good health. Non-Blacks were more likely than Blacks to quit, and more education favored cessation during the immediate postintervention period. Data from this study reported elsewhere indicate that the heaviest users of the manual were highly educated females.[10] These data also indicate that, although less-educated, Black females were consistent viewers of the televised intervention, they did not read the manual. Males were unlikely to have participated in any aspect of the intervention, but those who did were most likely to use the manual.

Thus, although these data are encouraging about the potential of such interventions, they need to be read with some caution. High and even moderate participation *were* correlated with higher rates of cessation than was low participation; in fact, those who registered and participated only slightly did worse at some waves than did the population panel, as can be seen from the bottom line in Table 3. On the other hand, this intervention may be reaching the same population that has responded to other interventions. This would imply that more attention must be given to developing materials that support the televised component and will make the intervention attractive to and used by those segments of the smoking population who are not being reached by current in-

TABLE 3—Cox Proportional Hazards Risk Ratios for Cessation across Waves of Interviews as a Function of Demographic Characteristics, Baseline Smoking Behavior, and Level of Participation in the Intervention

Variable	Postintervention	6 months	12 months	24 months
Gender				
Male	1.45* (1.16, 1.80)	1.21* (0.98, 1.50)	1.25* (1.08, 1.51)	1.19 (0.99, 1.43)
Race				
Black	0.79 (0.60, 1.03)	0.87 (0.67, 1.13)	0.83 (0.65, 1.06)	0.79* (0.63, 0.99)
Education				
College vs <high school	1.51* (1.00, 2.27)	0.89 (0.63, 1.25)	0.84 (0.63, 1.14)	0.94 (0.71, 1.26)
High school vs <high school	1.05 (0.68, 1.61)	0.80 (0.56, 1.14)	0.71* (0.52, 0.98)	0.83 (0.61, 1.13)
Health status				
Very good vs very poor	1.23* (1.04, 1.46)	1.24* (1.05, 1.47)	1.11 (0.96, 1.29)	1.15* (1.00, 1.32)
Average no. of cigarettes smoked daily				
30+ vs 0–19	0.57* (0.43, 0.75)	0.62* (0.47, 0.82)	0.53* (0.41, 0.68)	0.46* (0.36, 0.59)
20–29 vs 0–19	0.62* (0.48, 0.80)	0.78* (0.61, 0.99)	0.73* (0.58, 0.90)	0.71* (0.58, 0.88)
High participation vs population panel	7.45 (5.30, 10.49)	2.08* (1.45, 2.97)	1.27 (0.87, 1.86)	1.93* (1.38, 2.69)
Moderate participation vs population panel	3.15* (2.40, 4.13)	1.19 (0.92, 1.54)	1.11 (0.88, 1.42)	1.36* (1.08, 1.70)
Low participation vs population panel	1.54* (1.15, 2.07)	0.69* (0.53, 0.90)	0.78* (0.62, 0.99)	1.09 (0.88, 1.35)

*$P < .05$.

terventions and who are therefore likely to become the hard-core smoking population of the future.

Attrition was a problem; however, it is unlikely to explain differential outcomes because it was not related to baseline level of smoking or to smoking behavior at the time the subject was lost to follow-up.

Overall, the results of this paper encourage the use of an intervention in which television is combined with written, self-help materials. Taken with other results from this research,[10] they show a need for further investigation of the types of materials that are likely to be effective not only with those smokers who did not respond to the written materials used here but also in most other interventions of this kind. □

Acknowledgments

This research was supported by grant CA42760 from the National Cancer Institute.

The authors are especially grateful to the Chicago Lung Association, PruCare HMO, True Value Hardware, and particularly WLS-TV for their support of this research. Mary A. Spaeth of the Survey Research Laboratory, University of Illinois, is largely responsible for the clarity of the presentation although the authors are responsible for its content. We would also like to thank the two anonymous reviewers for their thoughtful comments on an earlier draft.

References

1. Thornberry OT, Wilson RE, Golden PM. Health promotion data for the 1990 objectives. Advance data from *Vital and Health Statistics,* No. 126. Hyattsville, Md: National Center for Health Statistics; 1986.
2. US Department of Health and Human Services, Public Health Service. *Promoting Health/Preventing Disease: Objectives for the Nation.* Washington, DC: DHHS; 1980.
3. Greenwald P, Sondik EJ. *Cancer Control Objectives for the Nation: 1985–2000.* Bethesda, Md: National Cancer Institute; 1986. NCI Monographs, National Institutes of Health publication 86–2880.
4. Pierce JP, Fiore MJ, Novotny TE, Hatziandreu EJ, Davis RM. Trends in cigarette smoking in the United States: projections to the year 2000. *JAMA.* 1989;261:61–65.
5. Pierce JP, Fiore MJ, Novotny TE, Hatziandreu EJ, Davis RM. Trends in cigarette smoking in the United States: educational differences are increasing. *JAMA.* 1989; 261:56–60.
6. Fiore MJ, Novotny TE, Pierce JP, Hatziandreu EJ, Patel KM, Davis RM. Trends in cigarette smoking in the United States: the changing influence of gender and race. *JAMA.* 1989;261:49–55.
7. US Department of Health and Human Services. *The Health Consequences of Smoking for Women: A Report of the Surgeon General.* Washington, DC: US Govt Printing Office; 1980.
8. Glynn TJ, Boyd GM, Gruman JC. Essential elements of self-help/minimal intervention strategies for smoking cessation. *Health Educ Q.* 1990;17:329–345.
9. Flay BR. *Selling the Smokeless Society.* Washington, DC: American Public Health Association; 1987.
10. Warnecke RB, Flay BR, Kviz FJ, et al. Characteristics of participants in a televised smoking cessation intervention. *Prev Med.* 1990;20:389–403.
11. Maccoby N, Farquhar JW. Communication for health: unselling heart disease. *J Commun.* 1975;25:114–126.
12. Ramirez AG, McAlister AL. Mass media campaign—*a Su Salud*. *Prev Med.* 1988; 17:608–621.
13. Warnecke RB, Langenberg P, Gruder CL, Flay BR, Jason LA. Factors in smoking cessation among participants in a televised intervention. *Prev Med.* 1989;18: 833–846.

Functional Limitations and Disability among Elders in the Framingham Study

ABSTRACT

Background. The measurement of physical disability as an indication of the impact of disease is commonly seen in research. However, these measures often do not clearly differentiate between functional limitations and daily performance of an activity.

Methods. We measured the differences between self-reported disability and observed functional limitations in six activities of daily living tasks among community-dwelling elders. The value of functional limitations vs disability measures in determining risk factors for disablement was ascertained.

Results. Systematic differences were found among the 1453 participants. At least 89% of the time when a difference was identified, the subjects ranked disability greater than the functional limitations observed. For those who were cognitively impaired, discrepancies occurred up to 11% of the time. In determining risk factors for disablement, we found that neurological impairments were associated with both functional limitations and disability, while sociocultural factors were associated with disability only.

Conclusions. Our findings suggest that physical functional limitations and disability in the elderly are two distinct concepts and that the measure of choice should be determined by research objectives and the type of population being studied. (*Am J Public Health.* 1992;82:841–845)

Margaret Kelly-Hayes, EdD, RN, Allen M. Jette, PhD, PT, Philip A. Wolf, MD, Ralph B. D'Agostino, PhD, and Patricia M. Odell, PhD

Introduction

Changes in life expectancy have resulted in an elderly population of increasing size and proportion where significant numbers are affected by disability.[1,2] Although physical limitations clearly increase with advancing age, reports of prevalence and severity of disability from existing studies may vary up to 3.1%. Although these differences are small in absolute terms, they can be large in percentage differences across surveys.[3] Methodological differences in obtaining the data and the type of population studied are two sources for these discrepancies.

Efforts to determine the magnitude of disability among elders have been hampered by lack of conceptual clarity.[4,5] Most of the current information on physical disability has been generated from self-report measures of activities of daily living (ADL). Often, these measures do not clearly differentiate between the presence of a functional impairment that makes an activity difficult or impossible to carry out and the actual performance of the activity.[6] Understanding of the methodology used to measure disability and determination of the specific functional limitations will not only improve precision for defining levels of disability among elders but also will increase our understanding of risk factors for disablement.

Important distinctions among concepts measurement in disability research have been addressed in a model developed by Saad Nagi[7,8] in which he proposes that there are three major consequences of disease or pathology. The first is impairment, representing the anatomical or physiological abnormality resulting from a pathologic process; the second, functional limitation, is the loss of ability to perform tasks and obligations of usual roles and normal daily life; and the third, disability, is an individual's behavior pattern that evolves with long-term impairment. In this report, we use Nagi's concept of functional limitation to describe a person's inability or restriction in performing an activity and disability to describe a person's pattern of behavior in daily life.

Documentation of the differences between a person's functional limitations and disability is sparse.[9] To distinguish these differences, we compared self-reported disability in six ADLs with observed functional limitations in the same activities in a controlled setting. Knowledge of the relationship between these two concepts and what factors may cause discrepancies will further understanding of disability measurement.

Methods

The study population was the Framingham cohort, a general population sample of residents of the town of Fram-

Margaret Kelly-Hayes is with the Department of Neurology, and Philip A. Wolf is with the Department of Neurology and the Section of Preventive Medicine and Epidemiology, Boston University School of Medicine, Boston, Mass. Allen M. Jette is with the New England Research Institute, Watertown, Mass. Ralph B. D'Agostino is with the Department of Mathematics, Boston University, Boston, Mass. Patricia M. Odell is with the Department of Mathematics, Bryant College, Smithfield, RI.

Requests for reprints should be sent to Margaret Kelly-Hayes, EdD, RN, Boston University School of Medicine, Department of Neurology, B603C, 80 East Concord Street, Boston, MA 02118.

This paper was submitted to the Journal September 27, 1990, and accepted with revisions August 22, 1991.

TABLE 1—Differences in Observed Functional Limitations and Self-Reported Disability in Activities in Daily Living (ADL)

ADL	Overall Differences		Self-reported Disability Greater than Observed Functional Limitations	
	%	n	%	No. of Differences
Dressing	3.9	1448	96.4*	56
Grooming	4.5	1453	96.8*	64
Feeding	3.1	1409	90.6*	43
Chair transfer	3.4	1446	97.9*	48
Walking	6.5	1448	96.8*	94
Stair climbing	6.5	1448	89.3*	94

*$P < .005$, extension of McNemar's test.

ingham, Mass. This cohort of 5209 men and women, aged 28 to 62 at entry in 1948, has routinely been assessed every 2 years for factors related to disease occurrence. The sampling procedures, clinical protocols, and disease determinants have been reported previously.[10,11] During a 2-year period (1983 to 1985), 1453 of the 1826 subjects participating in the 18th biennial cycle completed a disability questionnaire prior to their physical examination. At the time of this exam, 2758 of the original cohort of 5209 had died, 151 were institutionalized, and 474 had refused the biennial examination or were lost to follow-up.

Sixty percent of the participants (870 of the 1453) were women, the mean age was 72 years (range = 63 to 94), and all resided in the community. Because cohort members in nursing homes had on-site examinations and the controlled environment for testing functional limitations was not available, they were excluded from this analysis. The 373 cohort members taking Exam 18 who did not complete the disability questionnaire prior to the examination were also excluded. Those excluded were older (mean age = 76.7) and more likely to be female (64%).

Measures

Reported disability questionnaire. A questionnaire was mailed to Framingham participants 1 week prior to the examination. Each was asked to identify his or her normal day-to-day performance of six basic ADLs as defined by variables from the Katz and Mahoney and Barthel indexes.[12,13] These activities included dressing, grooming, feeding, transferring, walking, and stair climbing. Disability scores were assigned to each response (0 = "uses no help to perform the activity," 1 = "uses a device to perform the activity," 2 = "uses assistance of another person to perform the activity," and 3 = "does not perform the activity").

Functional limitation measure. At the beginning of the exam, a trained nurse measured each participant's ability to accomplish the specific tasks reported on the questionnaire according to a standard protocol. The functional limitation score ranged from 0 to 3 (0 = "did the activity without assistance," 1 = "did the activity using a device," 2 = "required human assistance to do the activity," and 3 = "was unable to do the activity").

Other variables. Several factors thought to influence disability measurement were included in the analyses. Age was dichotomized at below 75 years and 75 years and older. The Folstein Mini Mental State Exam (MMSE), a test of cognition, was administered.[14] MMSE scores range from 0 to 30, with scores of 23 or below reflecting probable cognitive deficits. Perceived health, whether a person lived alone or with others, and documentation of stroke or coronary heart disease, which included myocardial infarction, coronary insufficiency, and angina, were included.

Data Analysis

Cross tabulations between each functional limitation and disability item were performed to determine the frequency and direction of differences across each of the six activities. An extension of McNemar's test was used to determine the significance of differences in the discrepancies between functional limitations and disability,[15] and χ^2 tests were used to determine whether bivariate subgroup differences by age and cognition were statistically significant.

Multiple logistic regressions of discrepancies for each ADL were performed to determine the association with selected variables. The coding for the dependent variable for each ADL task was as follows: 0 = "agreement between functional limitations and self-reported disability in the ADL tasks" and 1 = "self-reported disability greater than functional limitations." The few cases in which the reported disability was less severe than the limitations observed were not included in the analyses. Odds ratios and 95% confidence intervals were calculated for each of the predictor variables. The independent variables were age, living situation, perceived health, cognitive function, stroke, and coronary heart disease.

To demonstrate the potential empirical impact of using functional limitations vs self-reported disability measures, we used two multivariate models to predict the presence of disability using both a sum of the observed functional limitation scores and a sum of the self-reported disability scores. For the first model, the sum of the functional limitations score was dichotomized as follows: 0 = "functionally independent in all ADL tasks by observation" and 1 = "functionally dependent in one or more ADLs by observation." The summary self-reported disability score was similarly dichotomized. The same predictor values were used as in the previous analyses.

Results

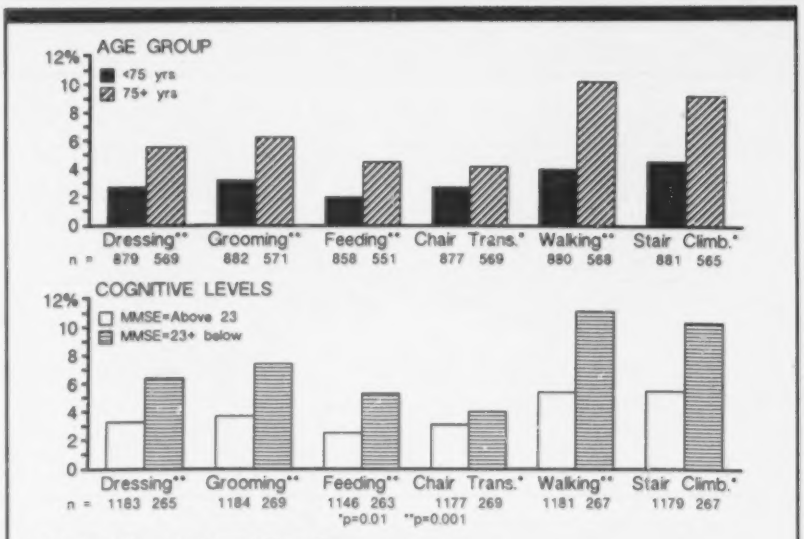
In this study of community-dwelling elders, over 92% of the cohort neither reported disability nor were observed to have functional limitations in the six ADL tasks. Differences did occur between functional limitations and disability in 3.1% to 6.6% of the cohort, most frequently in walking and stair climbing (6.5%) and least frequently in feeding (3.1%). The direction of the differences was highly systematic; at least 89% of the time when a difference was identified, the reported disability was greater than the functional limitations observed (Table 1).

Moreover, when the study sample was examined by age groups and cognitive levels, greater differences were identified. In all six ADL tasks, those who were older and those with cognitive impairment demonstrated significantly larger differences between reported disability and functional limitations (Figure 1). Disparities by age ranged from 3.2% to 4.7% in the below-75 age group and from 4.3% to 10.4% in those 75 years and older. Reported and observed limitations for mobility tasks showed the largest differences by age, with the older group having dif-

ferences of 10.4% in walking and 9.9% in stair climbing. When the sample was divided into those with low cognition scores (MMSE scores of 23 and below) and those within the normal range, the greatest differences between reported disability and observed limitations were seen, with the disparity between functional limitations and disability reaching 11.3% for the cognitively impaired. Again, the largest differences were observed in mobility activities.

The multiple logistic regressions (Table 2) revealed that age, cognitive impairment, and prior stroke were most strongly associated with disparities between functional limitations and reported disability. Stroke was significantly associated with differences in dressing, feeding, and stair climbing. Those in the older age group reported greater disability than the limitations observed in three of the six ADL tasks. Women, the cognitively impaired, those in the older age group, and those who perceived their health as fair or poor reported significantly more disability in walking, the ADL task with the most frequent differences.

To demonstrate the potential value of using a measure of functional limitations vs reported disability in determining risk factors for disablement, we calculated two multivariate models to predict physical disablement. Results of these summary analyses showed that only two of the seven predictors were significantly correlated with both functional limitations and disability scores (Table 3). The two strongest predictors of both functional limitations and disability were stroke and cognitive impairments. The three variables associated with reported disability, but not functional limitations, were age, sex, and perceived health.



Note. MMSE = Mini Mental State Exam.

FIGURE 1—Percentages by age groups and cognitive levels with disagreement between functional limitations and observed disability

Discussion

The results of this study demonstrate that in a cohort of community-dwelling elders, the vast majority neither reported disability nor were observed to have functional limitations. However, those with some restrictions/limitations rarely rated their disability below their functional limitations. These persons reported that their day-to-day disability in specific ADL tasks was greater than the observed functional limitations 89% of the time or more when a difference was identified. These results should dispel the persistent fear that self-report measures of functional disability will yield underestimates of the prevalence of disablement because of the older person's reluctance to admit to his or her limitations.[16] In fact, self-report of disability may yield higher estimates as compared with observed functional limitations if the self-report focuses on the individual's actual behavior in daily life.

Investigators are increasingly using performance testing in large-scale surveys of physical disability in older populations.[9] This trend may reflect legitimate concern that self-reported questions of physical capacities may elicit a subjective response that indicates what older persons think they are capable of, even though they may not normally attempt the activity. Such judgments are subject to potential error and may not accurately estimate the functional ability of the older person. A different challenge is faced when using self-reported disability measures. Here the question is whether or not a particular activity is not performed because of health problems or as a result of other unrelated reasons. Some investigators have used the concept of perceived difficulty in performing ADLs to distinguish the health-related significance of the reported behavior, while others have asked the respondents only to report limitations related to their health.[17-19]

This study examined the hypothesis that performance of ADLs has an integral cognitive component that influences day-to-day function. Those with cognitive impairment showed significantly greater differences between functional limitations and disability. This supports the need for further investigation into the specific manifestations and etiology involved in cognitive impairment and outcome measurements used in this population.[6] The fact that the older age group displayed significantly greater disparity for the ADLs studied, especially for mobility, may be the result of a more cautious life-style.

In general, mobility tasks more than self-care activities were associated with differences between functional limitations and disability. Disparities in walking were significantly associated with age, gender, cognitive impairment, and overall perceived health status of the cohort members.

The last goal of this study was to ascertain whether functional limitations and disability were associated with different risk factors. In other words, does the choice of one or the other influence the findings using the same predictive model? We found that cognitive impairment and history of stroke are the only two risk factors for functional limitations among the six variables in the logistic model. In contrast, disability was associated with four of the six factors, including age, gender, perceived health, and cognitive impairment and stroke. This supports Nagi's formulation that disability, unlike functional limitations, has a major social component.[7] Since disability reflects performance within a sociocultural context, one could expect that daily performance of an activity would be strongly influenced by sociocultural factors. One could speculate that

TABLE 2—Multiple Logistic Models Showing Association between Demographic Characteristics, Health Status, and Reported Disability Greater than Observed Functional Limitations

	Odds Ratios (95% Confidence Intervals)					
	Dressing	Grooming	Feeding	Chair Transfers	Walking	Stairs
Age[a]	1.77 (1.06–2.97)	1.76 (1.05–2.91)	1.55 (0.99–2.43)	1.23 (0.71–2.12)	1.75 (1.13–2.70)	1.23 (0.78–1.94)
Sex[b]	1.04 (0.62–1.74)	1.38 (0.82–2.33)	1.05 (0.67–1.64)	1.11 (0.65–1.91)	2.31 (1.41–3.77)	1.43 (0.92–2.28)
Living situation[c]	1.08 (0.72–1.64)	1.04 (0.69–1.55)	1.06 (0.73–1.53)	1.07 (0.68–1.67)	1.06 (0.75–1.49)	1.33 (0.95–1.87)
Perceived health[d]	1.05 (0.56–1.98)	1.30 (0.72–2.30)	1.04 (0.59–1.83)	1.97 (1.10–3.54)	2.79 (1.78–4.38)	2.30 (1.42–3.72)
Cognition (MMSE)[e]	1.99 (1.16–3.42)	1.63 (0.94–2.84)	1.00 (0.45–1.32)	1.00 (0.52–1.93)	1.62 (1.01–2.60)	1.87 (1.15–3.03)
Stroke[f]	2.70 (1.15–6.34)	1.44 (0.50–4.17)	2.31 (1.04–5.10)	1.16 (0.35–3.87)	1.71 (0.73–1.51)	2.29 (1.02–5.13)
Coronary heart disease[g]	1.11 (0.58–2.09)	1.22 (0.62–2.42)	1.23 (0.71–2.13)	1.22 (0.60–2.51)	1.60 (0.63–1.87)	1.23 (0.68–2.21)

Note. The dependent variable was coded as follows: 0 = "agreement between self-report and observed limitations" and 1 = "disagreement between self-report and observed limitations."
[a]1 = ≥75; 0 = <75.
[b]1 = female; 0 = male.
[c]1 = living with others; 0 = living alone.
[d]1 = fair or poor; 0 = good or excellent.
[e]MMSE = Mini Mental State Exam. 1 = MMSE ≤ 23 (impaired); 0 = MMSE > 23 (normal).
[f]1 = yes; 0 = no.
[g]1 = yes; 0 = no.

TABLE 3—Logistic Regression Models of Predictors of Obeserved Functional Limitations and Disability for All ADL Tasks

	Dependent Variable			
	Functional[a] Limitations		Disability[b]	
	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval
Age (<75/75+)	0.72	0.45–1.17	2.1	1.41–3.10
Sex (male/female)	0.84	0.53–1.33	1.52	1.01–2.28
Living situation (alone/with others)	0.91	0.63–1.33	0.82	0.61–1.12
Perceived health (excellent or good/fair or poor)	1.32	0.97–1.80	2.03	1.57–2.62
Cognition (≤23/>23 on MMSE)	2.55	1.27–5.12	2.58	1.41–4.71
Stroke (absent/present)	3.13	1.47–6.65	2.56	1.27–5.19
Coronary heart disease (absent/present)	0.64	0.31–1.35	1.0	0.49–2.03

Note. MMSE = Mini Mental State Exam.
[a]0 = no functional limitations in all six ADL tasks, 1 = functional limitations in one or more ADL tasks.
[b]0 = no disability in all six ADL tasks, 1 = disability in one or more ADL tasks.

these factors are proxies for other variables that influence disability but not actual functional limitations.

Limitations in the design of this investigation prevent us from determining conclusively the extent to which differences were due entirely to the difference between the functional limitations and reported disability dimensions being assessed. The disability measures differed with respect to the methods used to assess the older person's function: the functional limitations measure relied upon the professional's report of observed behavior, while the reported disability measure relied upon the self-report of the elder. Our analyses are unable to differentiate the degree to which observed differences are due to these different methods. Factors that may influence reported disability include mood, pain, and general feeling of health at the time the questionnaire is completed. Also, the estimates of disability are limited to those who were able to be evaluated at the clinic site, where observation by a professional could influence performance. Therefore, they do not represent those who are more likely to be more disabled at home or in an institution.

Our analyses do not suggest a clear preference for using either functional limitations or disability as the most appropriate outcome variable. Rather, we recommend that the choice between the two measurements be based on the objective of the specific research hypotheses and the population under study. For descriptive questions, this analysis suggests that estimates of prevalence will be similar regardless of whether the focus is on functional limitations or disability. Greater concern, however, should be exercised when the research focus is on high-risk subgroups such as the cognitively impaired, where prevalence estimates can differ to a greater extent. Although differences in prevalence estimates may be small, the choice of measuring functional limitations or disability may seriously affect empirical results. Risk factors for disability and functional limitations may be quite different, as posited by Nagi over 25 years ago. □

Acknowledgments

This research was supported in part by a contract from the National Institute of Health/National Heart, Lung, and Blood Institute (No. 1-HC-38038) and Grant 2-RO1-NS-17950 (National Institute of Neurological Disorders and Stroke).

References

1. National Center for Disease Statistics. *Current Estimates from the National Health Interview Survey, 1988.* Washington, DC: US Government Printing Office; 1989. DHHS publication no. (PHS) 89-150. (Vital and health statistics; series 10; no. 173).
2. National Center for Health Statistics. *The National Nursing Home Survey, 1985 Summary for the United States.* Washington, DC: US Government Printing Office; 1989. DHHS publication no. (PHS) 89-1758. (Vital and health statistics; series 13; no. 97).
3. Weiner JM, Hanley RJ, Clark R, Van Nostrand J. Measuring the activities of daily living among the elderly: a guide to national surveys. *J Gerontol.* 1990;45:S229–S287.

4. Pope A, Taylor A, eds. *Disability in America: Toward a National Agenda for Prevention.* Washington, DC: National Academy Press; 1991.
5. McDowell I, Newell C. *Measuring Health: A Guide to Rating Scales and Questionnaires.* New York, NY: Oxford University Press; 1987.
6. Jette AM, Branch LG. The Framingham disability study: II. physical disability among the aging. *Am J Public Health.* 1981;71:1211–1216.
7. Nagi SZ. Some conceptual issues in disability and rehabilitation. In: Sussman MB, ed. *Sociology and Rehabilitation.* Washington, DC: American Sociological Association; 1965:100–113.
8. Nagi S. Disability concepts revisited: implications for prevention. In: Pope A, Taylor A, eds. *Disability in America: Toward a National Agenda for Prevention.* Washington, DC: National Academy Press; 1991.
9. Guralnik JM, Branch LG, Cummings SR, Curb JD. Physical performance measures in aging research. *J Gerontol.* 1989;44: M141–M146.
10. Gordon T, Moore FE, Shurtleff D, et al. Some methodological problems in the long term study of cardiovascular disease: observations from the Framingham Study. *J Chronic Dis.* 1959;10:186–206.
11. Shurtleff D. *Some Characteristics Related to the Incidence of Cardiovascular Disease and Death: The Framingham Study—18 Years Follow-up.* Washington, DC: US Department of Health, Education, and Welfare; 1984. Section 30, publication no. (NIH) 74–599.
12. Katz S, Downs TD, Cash HR, et al. Progress in the development of an index of ADL. *Gerontologist.* 1979;10:20–30.
13. Mahoney FJ, Barthel DW. Functional evaluation. the Barthel Index. *Md Med J.* 1965;14:61–65.
14. Folstein MF, Folstein S, McHugh PR. Mini Mental State Exam: a practical method for grading the cognitive state of patients for clinicians. *J Psychiatr. Res.* 1975;12:189–198.
15. Dixon WJ, Massey FJ. *Introduction to Statistical Analysis.* New York, NY: McGraw-Hill; 1983:288–291.
16. Miniken KL, Rowe JW. Health and disease among the oldest old: a clinical perspective. *Milbank Q.* 1985;63(2):324–349.
17. Patrick DL, Dorby SC, Green S, Horton G, Locker D, Wiggins R. Screening for disability in the inner-city. *J Epidemiol Community Health.* 1981;35:65–70.
18. Jette AM, Dennison DL. Interobserver reliability of a functional status assessment instrument. *J Chronic Dis.* 1978;311:503–513.
19. Verbrugge L. The iceberg of disability. In Stahl S, ed. *The Legacy of Longevity: Health and Health Care in Late Life.* Newbury Park, Calif: Sage; 1990.

Twelfth International Conference on Alcohol, Drugs, and Traffic Safety

The 12th International Conference on Alcohol, Drugs, and Traffic Safety will be held in Cologne, Germany, September 28 to October 2, 1992, and will consider the question of which blood alcohol limits to set for drivers in the united Germany and in the future united Europe. Traffic safety experts from all over the world, legal and medical experts, and psychologists will aim to contribute to a worldwide reduction in the traffic deaths and injuries attributable to drug and alcohol abuse.

The conference will be comparing the effects of legal alcohol limits and police surveillance tactics in different countries. It will also look at alcohol consumption trends in various cultures, the effects of alcohol on driving behavior and accident occurrence, the influence of medication and drugs, and prevention and rehabilitation.

The International Committee for Alcohol, Drugs, and Traffic Safety, formed in Stockholm in 1950, holds a conference every three years on alcohol and driving. The Cologne Conference is being sponsored by the German Federal Ministry of Transport and organized and supported by the TÜV-Akademie Rheinland and the Medical and Psychological Institutes of TÜV Rheinland and TÜV Hannover.

If you register before June 30, 1992, the fee for the five-day conference is DM 750 (including taxes). For further information or to register, contact the conference office at TÜV-Akademie Rheinland, Am Grauen Stein, 5000 Cologne 91, Germany; tel. 0221 806-3062, fax 0221 806-3006.

Neuroleptic Medication and Prescription Practices with Sheltered-Care Residents: A 12-Year Perspective

Steven P. Segal, PhD, David Cohen, PhD, and Stephen R. Marder, MD

ABSTRACT

Objectives. Most adult residents of sheltered-care facilities (board and care, family care, psychosocial rehabilitation, and other supported housing arrangements) for the chronically mentally ill receive neuroleptics. These facilities house over 300 000 mentally ill residents, but neuroleptic prescription practices with this population have not been studied.

Methods. A probability sample (n = 393) of all adult former psychiatric patients in sheltered care in California was surveyed in 1973; 94% of the located survivors (n = 243) were reinterviewed 12 years later.

Results. In 1973, 79% received neuroleptics; in 1985, 76%. Polypharmacy decreased, and the elderly remained less medicated than adults. Yet, mean daily neuroleptic doses doubled, more persons received higher doses, and 62% reported adverse effects. Furthermore, high dosing was attributed to psychiatrists rather than other physicians, even when controlling for residents' clinical and sociodemographic characteristics.

Conclusions. Neuroleptic drugs became the staple pharmacological treatment for mentally ill sheltered-care residents. While physicians more cautiously medicated the elderly, they had not reduced doses by 1985, even after a decade of treatment. The specialty of the prescriber was an important factor in preference for high-dose treatment. (*Am J Public Health.* 1992;82:846–852)

Introduction

More than three quarters of the adult residents of sheltered-care facilities with a history of mental hospitalization are prescribed neuroleptic drugs.[1] Because long-term use of neuroleptics, which delays relapses in over 50% of schizophrenic patients,[2] also poses great risk of persistent tardive dyskinesia,[3] researchers and professional associations have recently urged clinicians to exercise caution when prescribing these drugs.[4,5] This study describes psychiatric drug prescriptions in a cohort of residents in sheltered care (board and care, family care, halfway houses, and psychosocial rehabilitation facilities). Specifically, we consider prescription practices in the mid-1980s to determine how they changed from the early 1970s.

A partial consensus concerning the use of neuroleptics emerged in the 1980s around four issues. First, although medications such as antiparkinsonians or antidepressants are commonly added to neuroleptic regimens, the simultaneous prescription of more than one neuroleptic or more than three psychotropics has been particularly criticized.[6,7] Second, there has been considerable interest in low-dose strategies,[8–11] which are sometimes associated with higher relapse rates but usually with improved social functioning and fewer adverse effects.[12–15] Third, although high- and low-potency neuroleptics are considered equally effective,[16,17] maintenance treatment with injected depot preparations has been recommended[10,14,18] to circumvent most patient noncompliance and ensure exact dosage delivery. Fourth, since the likelihood of noncompliance, adverse effects, and errors in administration increases with the frequency of administration, once-a-day dosing has been recommended for oral maintenance treatment.[19–21]

Most experts writing in the mid-1980s recommended general adherence to these four prescription practices, but the subject has not been studied extensively in recent years and the literature abounds with contradictory anecdotal impressions.[22,23] No study has examined prescription practices in a cohort of former patients living in community-based sheltered-care settings. This is an important omission, for these facilities have become the placement of choice in the mental health system, housing 300 000 to 400 000 mentally ill adults.[24]

Methods

This 12-year longitudinal study first involved face-to-face, structured interviews in 1973 with 393 residents of 211 sheltered-care facilities in 157 census tracts in California, a self-weighting probability sample of all 18- to 65-year-old formerly hospitalized ex-patients (n = 12 430) in sheltered care in the state.

The sample was obtained by dividing the state into three master strata: Los Angeles County, the nine-county San Francisco Bay Area, and all other counties. In the first two, facilities were stratified by

Steven P. Segal is with the University of California, Berkeley. David Cohen is with the Université de Montréal, Montreal, Quebec, Canada. Stephen R. Marder is with the University of California, Los Angeles.

Requests for reprints should be sent to Steven P. Segal, PhD, Professor and Director of Mental Health and Social Welfare Research Group, School of Social Welfare, 120 Haviland Hall, University of California, Berkeley, CA 94720.

This paper was submitted to the Journal June 10, 1991, and accepted with revisions November 27, 1991.

size and a sample was drawn of facilities with probabilities proportionate to bed capacity. In the third stratum, a cluster sample was designed, selecting two counties from the northern and two from the southern part of the state with probability proportionate to size. From each pair, samples of facilities were selected, also with probability proportionate to size. Residents were sampled within facilities, using systematic random sampling from specially prepared field listings.[25]

Follow-up data were collected in 1985. Of the 393 former residents, 360 (91.6%) were located and 33 (8.4%) were not. Of the former residents located, 270 (68.7%) were alive and 90 (22.9%) confirmed dead. Of the survivors, 253 (93.7%) were interviewed, and 243 (90%) yielded complete drug information. This study focuses on psychotropic drug prescriptions to these 243 survivors (61% of the original sample). Table 1 shows no differences between the two samples except for the survivor sample being younger and having more women and fewer Blacks than the original, due to higher death rates among the elderly, men, and Blacks.

In 1973 and 1985, drug prescription data (name of drug, dosage, times per day, number and frequency of injections) were obtained from respondents and their residence managers or physicians. All active prescriptions at the time of the interview were recorded. Self-reported dosages were confirmed by checking prescription bottles. Neuroleptic dosages from both years were converted into oral chlorpromazine equivalents (CPZeq) using Hollister's[26] scale. For the injectable, long-acting fluphenazine decanoate, CPZeq was determined by a first equivalence to daily units of fluphenazine hydrochloride, using the rule endorsed by the manufacturer of Prolixin® and by other psychopharmacologists[27,28]: 12.5 mg of fluphenazine decanoate injected every 3 weeks is equivalent to 10 mg of fluphenazine hydrochloride daily. Unless otherwise noted, all dosages are reported in mg/day CPZeq.

Information on adverse drug effects was obtained only during the 1985 study, with respondents asked to rate the severity of 13 effects from "absent" to "very severe."

Chi-square tests, two-tailed *t* tests, and multiple regression analysis were performed on the data using the SPSSx statistical package. Significant results are reported when $P < .05$.

TABLE 1—Characteristics of Chronic Mentally Ill Patients (n = 393) Living in Sheltered-Care Facilities in California in 1973, and Cohort Survivors with Psychotropic Drug Data (n = 243)

Characteristic	Total, %	Survivor Sample, %
Sex		
Male	57.0	53.5*
Female	43.0	46.5
Age, y		
18 to 35	27.0	32.9*
36 to 45	15.5	17.7
46 to 55	30.3	30.5
56 to 65	27.2	18.9
Ethnic group		
Black	13.0	10.0*
White	76.0	77.6
Other	11.0	12.4
Education, y		
0–6	9.6	9.2
7–12	66.8	67.6
13 or more	23.6	23.1
Marital status		
Single	58.1	60.9
Married	4.9	5.3
Formerly married	37.1	33.7
Employment		
Works for cash or other benefits		
Never	73.6	71.0
At least rarely	26.4	29.0
Months in current sheltered-care facility		
1	5.6	5.2
2 to 12	35.8	36.9
13 to 24	18.0	17.2
25 or more	40.6	40.8
Has spent a continuous period of 2 or more years in a state psychiatric hospital (pre-1973)		
Yes	40.5	43.5
No	59.5	56.5
Psychiatric admission in the past year (pre-1973)		
Yes	21.7	23.9
No	78.3	76.1

*Significantly different ($P < .05$) from the nonsurviving group.

Results

Cohort Status

In 1973, 213 persons (87.6%) received a psychotropic drug prescription, compared with 206 (85.2%) 12 years later.

In 1973, 79% of the residents were prescribed neuroleptics, 81% of women and 77% of men. By 1985, 76% received these drugs, and women were significantly more likely to be prescribed neuroleptics (82% vs 71%, $\chi^2 = 3.8$, $df = 1$, $P < .05$), the only significant gender difference detected in the prescription of any psychotropic. Table 2 shows how many patients were prescribed how many drugs per class.

Diagnosis and medication. Lifetime modal diagnoses based on hospital discharge data were determined for 209 of the 243 sample members. Most (79%) received diagnoses of schizophrenic disorders, the rest substance abuse (5%) or other disorders (16%). Modal diagnoses for 162 (88%) of those on neuroleptics in 1985 were: schizophrenia or other psychoses (87%), affective disorder (6%), and other disorders (7%).

Diagnoses were available for 26 (93%) of 28 persons prescribed an antidepressant in 1985. Only 5 (20%) had any diagnosis of affective disorder noted in their lifetime hospital records. Of 9 persons prescribed lithium at follow-up, only 2 had any such diagnosis. There were no sex or age differences in the likelihood of receiving these drugs without the diagnosis.

Neuroleptic drugs. Ten different neuroleptics were prescribed in 1973, eight at follow-up. Of 192 persons on neuroleptics in 1973, 164 (85%) still received them 12 years later. In the interval, only thio-

TABLE 2—Number of Drugs in Each Class and Number and Percentage of Persons Receiving Psychotropic Drugs, by Drug Class

Drug Class	1973			1985		
	No. of Drugs in Class	No. Using Each Class	% Using Each Class	No. of Drugs in Class	No. Using Each Class	% Using Each Class
Neuroleptic	10[a]	192	79.0	8[a]	185	76.1
Antiparkinsonian	5	98	40.3	7	90	37.0
Anticonvulsant	2	26	10.7	3	16	6.6
Sedative–hypnotic	8	20	8.2	2	10	4.1
Antidepressant	4	15	6.2	7	28	11.5
Anxiolytic	2	10	4.1	2	7	2.9
Lithium	1	2	0.8	1	9	3.7
Antisenility	0	0	...	1	3	1.2
Stimulant	0	0	...	1	1	0.4

[a]Fluphenazine hydrochloride and fluphenazine decanoate are counted as a single drug.

TABLE 3—Patients Receiving Each of the Six Most Popular Neuroleptics (n = 243)

Neuroleptic	1973			1985		
	%[a]	Range	MDD ± SD	%[a]	Range	MDD ± SD
Thioridazine	35	20–900	254 ± 185	22	20–800	192 ± 163[b]
Trifluoperazine	28	4–100	20 ± 16	17	2–200	25 ± 36
Chlorpromazine	25.5	15–1200	276 ± 239	11.5	30–2250	392 ± 440
Fluphenazine hydrochloride	8	1–20	8.5 ± 6	14	2–200	38 ± 42
Thiothixene	4	6–60	22 ± 16	6	4–40	11.5 ± 10
Fluphenazine decanoate	2.5	750–1500[c]	375 ± 306[c]	7	300–4500[c]	2040 ± 1171[c]
Haloperidol	2	1–60	20 ± 23	11.5	1–50	19 ± 14
Others[d]	4	...	...	2.8	...	...

Note. MDD = mean daily dose.
[a]Total exceeds 100% as patients may have received more than one drug.
[b]$t = 1.97$, $df = 30$, $P < .05$.
[c]In mg/day chlorpromazine equivalent.
[d]In 1973, these are piperacetazine, prochlorperazine, reserpine, and acetophenazine. In 1985, they are chlorprothixene and perphenazine.

ridazine's mean daily dose (MDD) showed a significant reduction (see Table 3).

Prescription Practices

Polypharmacy. In 1973, the mean (±SD) number of drugs prescribed per person was 1.9 (±1.2), reduced to 1.7 (±1.2) in 1985 ($t = 2.13$, $df = 242$, $P = .03$). However, the proportion receiving three psychotropics dropped from 21.4% to 13.2%. Over the study interval, 34% received the same number, 38% received fewer drugs, and 28% received more drugs. Table 4 highlights a reduction of multiple neuroleptic prescriptions, from 29% of the whole sample prescribed two or more neuroleptics in 1973 to 13% by 1985.

Among the medicated elderly, 51% were prescribed one drug, 27% two drugs, and 23% three or more. The comparable percentages for the younger group are 31%, 44%, and 25%, respectively ($\chi^2 = 8.83$, $df = 2$, $P = .012$).

Neuroleptic drug dosage. Among those prescribed neuroleptics in both 1973 and 1985, the MDD (±SD) increased from 512 (±435) mg in 1973 to 963 (±1603) mg in 1985 ($t = 3.64$, $df = 163$, $P = .0001$). Six percent were prescribed the same dose both years, and the rest were evenly split between decreases (46%) and increases (48%). After excluding fluphenazine decanoate, a significant MDD increase was still observed, from 461 (±371) mg to 790 (±1480) mg ($t = -2.72$, $df = 143$, $P = .007$).

The range of neuroleptic dosages increased from 30 to 3400 mg in 1973 to 20 to 11 300 mg in 1985. While only 4 persons (1.6%) in 1973 were prescribed doses over 2000 mg daily, this had increased to 20 persons (9%) by 1985, of which only 3 resided in institutions (where dosages are typically higher). The proportion receiving over 1000 mg daily doubled over the study interval, from 11% to 23%.

Long- and short-acting injectables. In 1973, 6 people (2.5%) received neuroleptics by injection, all of them of fluphenazine decanoate. At follow-up, 22 people (9%) received injections, 16 (7%) of the decanoate products and the others of short-acting neuroleptics.

The MDD of those receiving injections in 1985 was four times greater than that of those receiving oral prescriptions: 2691 (±2585) mg vs 662 (±1130) mg ($t = 3.73$, $df = 21$, $P = .0001$). The mean age in the first group was 46 (±11) years, compared with 55 (±12) years for others ($t = 3.26$, $df = 182$, $P = .001$). They were also prescribed an average of 3 (±1.3) psychotropic drugs, compared with 1.9 (±0.9) for those on oral neuroleptics ($t = -3.70$, $df = 24.16$, $P = .001$).

Age, gender, and dosage. In 1973, the MDD (±SD) among the neuroleptized elderly was 308 (±424) mg (range: 50 to 2060 mg). By follow-up, this had increased to 484 (±753) mg (range: 25 to 4500 mg) but was still less than half that prescribed to persons under 60 (1154 [±1793] mg). This elderly-to-adult dose ratio follows

well-established guidelines to prescribe one third to one half the usual adult dose.[29] However, whereas 36% of elderly on neuroleptics in 1973 were receiving low (under 150 mg) daily doses and 27% were receiving high doses (over 300 mg), by 1985 these proportions were reversed. Over the study interval, increases in MDDs were significant among both men and women, but changes in dosage were not related to gender.

High- and low-potency neuroleptic. Of those prescribed neuroleptics both years, 40% remained on the same potency at follow-up, 18% switched from low to high, and only 5% went from high to low. Most who started neuroleptic treatment after 1973 were prescribed high-potency drugs at follow-up.

During both study periods, high-potency agents were prescribed at significantly higher CPZeq doses than low-potency drugs, and this relative difference had greatly increased by 1985. The high-potency to low-potency dose ratio went from 1.74:1 in 1973 to 5.3:1 in 1985 (excluding fluphenazine decanoate at follow-up, the ratio was 4.6:1).

At follow-up, people receiving neuroleptics of both potencies (n = 25) were (1) more likely to be prescribed antiparkinsonians (56% of them and 52% of the high-potency group, compared with 32% of the low-potency group; $\chi^2 = 7.16$, $df = 2, P = .03$); (2) prescribed more psychotropic drugs on average (3.2 [±0.9], compared with 1.9 [±1.1] for those on high-potency drugs and 1.6 [±0.8] for those on low-potency drugs [F = 21.4, $P = .0001$]); and (3) prescribed drugs more frequently (2.9 [±1.1] times per day, compared with 2.4 [±0.9] for those on high-potency drugs and 2.1 [±1] for those on low-potency drugs [F = 3.9, $P = .02$]). No significant age or sex differences between the three groups were noted.

Daily dosing schedule. The mean maximum daily dosing schedule decreased nonsignificantly from 2.2 (±1.3) times in 1973 to 2.0 (±1.3) times in 1985.

TABLE 4—Psychotropic Drug Combinations

Drug Combination	Prescriptions 1973 n	1973 %	1985 n	1985 %
No neuroleptic	51	21.0	58	23.9
Neuroleptic × 1	121	49.8	154	63.4
Neuroleptic × 2	66	27.2	27	11.1
Neuroleptic × 4	5	2.0	4	1.6
Total	243	100.0	243	100.0
No psychotropic drugs	30	12.3	37	15.2
Neuroleptic × 1 only	78	32.1	81	33.7
Neuroleptic(s) + antiparkinsonian only	68	28.0	61	25.1
Neuroleptic(s) + antiparkinsonian + 1 other psychotropic	21	8.6	17	7.0
Neuroleptic(s) + lithium only	0	0.0	5	2.1
Neuroleptic(s) + anticonvulsant only	6	2.5	2	0.8
Neuroleptic(s) + sedative–hypnotic only	5	2.1	1	0.4
Neuroleptic(s) + other drug combinations	14	5.8	17	7.0
Other drug combinations	21	8.6	22	9.1
Total	243	100.0	243	100.0

Practice Differences between Psychiatrists and Other Physicians

At follow-up, the medical specialty of the current prescribers of 115 (73%) of the patients on neuroleptics in 1973 and 1985 was known. For people (n = 84) with prescriptions from psychiatrists, the MDD (±SD) increased from 483 (±440) to 1087 (±1648) mg ($t = -3.46$, $df = 83$, $P = .001$), while for those (n = 31) with prescriptions from nonpsychiatric physicians, the MDD (±SD) decreased from 1973 to 1985 from 431.5 (±377) to 263 (±295.5) mg ($t = 2.06$, $df = 30$, $P = .048$). No difference existed between dose levels of these two groups in 1973. This same pattern was observed for the 48 people over age 59 in 1985. For the 20 elderly with prescriptions from other physicians in 1985, the MDD (±SD) decreased from 495 (±447) to 202 (±196) mg ($t = 2.88$, $df = 19$, $P = .009$). For the 28 with prescriptions from psychiatrists, the MDD (±SD) increased from 339 (±231) to 667.5 (±780) ($t = -2.24$, $df = 27$, $P = .033$).

More problematic patients may be referred primarily to psychiatrists, so we sought to determine whether their patients differed from those seen by other physicians. We found no significant differences between them in total time spent in the hospital, number of hospitalizations between 1973 and 1985, Brief Psychiatric Rating Scale (BPRS) scores in 1985, or type of residence in 1985. However, those with prescriptions from psychiatrists were younger (mean age of 52.55 years vs 58.31 years; $t = 2.86$, $df = 155$, $P = .005$).

To predict CPZeq dose at follow-up, we entered the following variables in a multiple regression: age, 1973 BPRS score, 1985 BPRS score, 1973 CPZeq dose, number of hospitalizations, location in an institutional setting in 1985, and whether the neuroleptic was prescribed by a psychiatrist or other physician in 1985. Our model explained 22% of the variance in 1985 CPZeq dose (F = 5.87, $df = 7.148$, $P = .0000$). Significant predictors were the number of hospitalizations ($t = 2.06$, $P = .04$), age ($t = -1.99$, $P = .05$), and who prescribed the neuroleptic ($t = 2.46$, $P = .01$). Psychiatrists were more likely to prescribe at higher doses, even after taking all other factors into account.

We also examined people (n = 41) who were receiving over 1000 mg/day CPZeq, considered excessive by American Medical Association standards.[30] Of these 41, the specialty of the prescriber was available for 27 people. Of these 27, 26 (96%) were prescribed their neuroleptic by a psychiatrist. Mean age among those 26 was 47.5 years and mean time since their last hospitalization was 7.2 years.

Most neuroleptic dose-level differences observed here result from a psychiatric preference for high-potency drugs. Of those people with neuroleptic prescriptions from psychiatrists, 72% received high-potency drugs, compared with 46% of those who saw another type of physician ($\chi^2 = 6.82$, $df = 2$, $P = .0329$).

Adverse Effects

Of the 185 people on neuroleptics in 1985, 62.5% reported at least a mild degree of 1 or more of 13 adverse effects; 26.7% reported experiencing 1 to 2 effects, and 35.8% reported 3 to 13 effects (see Table 5).

Age, gender, and adverse effects. For each adverse effect except weight gain/loss, a greater proportion of people under 40 years reported discomfort. Dry mouth/throat, inability to sit still, depression/apathy, skin rashes, and loss of sex drive were reported significantly more by

TABLE 5—Prevalence of Self-reported Medication Side Effects, 1985

Side Effect	Degree of Severity: Mild n	Mild %	Moderate n	Moderate %	Severe or Very Severe n	Severe or Very Severe %	Total %
Dry mouth or throat (n = 175)	27	14.6	21	11.4	19	10.2	36.2
Restlessness/can't sit still (n = 176)	19	10.3	21	11.4	10	5.4	27.1
Weight gain/loss (n = 172)	18	9.7	14	7.6	9	4.9	22.2
Diarrhea/constipation (n = 175)	21	11.4	8	4.3	12	6.5	22.2
Depression/apathy (n = 173)	18	9.7	11	5.9	6	3.2	18.8
Dizziness/weakness (n = 175)	18	9.7	8	4.3	9	4.8	18.8
Skin rash/sensitive to sun (n = 175)	12	6.5	15	8.1	5	2.7	17.3
Problems staying awake (n = 178)	15	8.1	8	4.3	4	2.1	14.5
Sight/eye problem (n = 173)	8	4.3	9	4.9	9	4.9	14.1
Loss of sex drive (n = 171)	4	2.2	10	5.4	10	5.4	13.0
Body rigidity/stiffness (n = 175)	9	4.9	4	2.2	8	4.3	11.4
Spasms of head/neck/mouth/jaw (n = 176)	4	2.2	7	3.8	3	1.6	7.6
Nausea/vomiting (n = 175)	9	4.9	3	1.6	2	1.1	7.6

younger patients. Women and men did not differ in the number of side effects they reported, but women were significantly more likely to report dizziness/weakness, changes in weight, and sight or eye problems.

Adverse effects and prescription practices. Scores of the severity ratings of all 13 adverse effects were summed to produce a total score (possible range: 0 to 52) ranging from 0 to 30 (mean [±SD] = 4.56 [±0.1]). Most people (38%) scored 0; 27% scored from 1 to 4, 20% from 5 to 12, and 15% from 13 to 30. Exactly half the sample reported two mild or one moderately severe side effect. Both the number ($r = .17$, $P < .01$) and the severity ($r = .17$, $P < .01$) of adverse effects showed modest positive correlations with neuroleptic dosage but no significant relationship with neuroleptic potency.

Discussion

Between 1973 and 1985, some psychotropic prescription practices with sheltered-care residents were modified while others remained unchanged. On the positive side, nearly 40% of the sample received fewer drugs at follow-up. Multiple neuroleptic prescriptions were reduced by more than half. The proportion of people prescribed neuroleptics and antiparkinsonians stayed the same over the 12-year period, confirming that these drugs constitute the staple pharmacological treatment for long-term patients. Similarly, the elderly continued to receive, on average, half the neuroleptic dosage prescribed adults. However, overall changes in neuroleptic dosage patterns contrast with published statements such as "unquestionably, concern about tardive dyskinesia has led to more conservative use of antipsychotic drugs; i.e., more careful selection of patients for long-term therapy and the use of lower doses."[22]

We found that 85% of those prescribed neuroleptics in 1973 received them 12 years later and observed substantially higher mean daily doses across age, sex, and ethnic categories. Nearly half of those prescribed neuroleptics both years experienced a dose increase in the interval. Nearly 10% received over 2000 mg CPZeq daily, an extreme dose usually accompanied by polypharmacy and more frequent drug administration. Reflecting their current acceptance as specific treatments for affective disorders, antidepressants and lithium were increasingly prescribed, but most commonly as adjuncts to neuroleptics.[21]

The large increase in MDD is associated with growing use of high-potency neuroleptics. This longitudinal study of long-term sheltered-care patients confirms reports from cross-sectional studies that higher CPZeq doses are prescribed when using high- vs low-potency neuroleptics.[31,32] In this sample, the high- to low-potency ratio is two times that reported in previous studies. This should be of concern, since some evidence suggests that, compared with moderate doses of neuroleptics, higher doses of high-potency drugs frequently produce neurotoxicity and inferior antipsychotic effect.[30,33]

These patterns of dissimilar dosing might be explained by the fact that low-potency agents carry an increased risk of inducing autonomic dysfunction and systemic toxic reactions that are less common or severe when lower doses are used.[34] Some authors[33] find good evidence that high CPZeq doses are not more beneficial than moderate doses but believe that "the recently popular aggressive use of high-potency neuroleptic agents in high doses is to increase the degree and speed of therapeutic response." Our findings indeed suggest that prescribers (primarily psychiatrists) are opting for drugs permitting large dose increases. The higher incidence of extrapyramidal effects associated with these drugs may be considered more acceptable than the autonomic and other physical effects associated with high doses of low-potency drugs.

Patients who had medication switched from low potency to high potency may have been poor responders initially who showed improvement with a drug change and a dose increase. On the other hand, studies indicate that only a small proportion of poor responders at conventional dosages will respond at substantially higher doses.[35,36] That nearly 40% were prescribed the same neuroleptic 12 years apart may denote a stable clinical profile and the patient's acceptance of drug treatment. However, if clinicians should aim to decrease total exposure to neuroleptics, that 60% were switched is disturbing, since nearly two thirds who were switched received a higher CPZeq dose. A possible development of tolerance to neuroleptics could explain increased doses over time in the same individuals, especially if treating physicians remained the same, but we possessed no data on this variable. Possibly, increase in MDD with high-potency neuroleptic was a response to observations in

the late 1970s supporting high doses and rapid neuroleptization in hospital settings,[37] a practice currently in relative disfavor. Patients may have been discharged on higher doses that, in turn, increased during the next hospitalization or convinced clinicians that higher doses were relatively safe. We expect that these attitudes changed during the late 1980s, when concerns about toxic effects of potent neuroleptics were raised[38]: the continuation of our follow-up study to include 20-year evaluations in 1993 will either support or disprove this interpretation.

Nevertheless, in both 1973 and 1985, 4 in 10 patients received neuroleptic doses under 300 mg CPZeq per day, confirming that many chronic schizophrenic patients in the community can be maintained in lower dose ranges.[13,39,40] On the other hand, low-dose strategies must be complemented by community services that can respond effectively to signs of impending relapses. Unfortunately, the level of services available for this population may make low doses appear impractical to prescribers.

One subset of prescribers, nonpsychiatric physicians, appears to follow more cautious guidelines than their psychiatric colleagues, perhaps because of the former's lesser familiarity with neuroleptics and managing adverse effects. We could not locate similar observations in the literature; these practice differences require further investigation.

Nearly 40% of our respondents reported moderate to very severe adverse effects. Restlessness, body rigidity, or orofacial spasms—effects indicative of extrapyramidal symptoms—were reported by 33% of the sample. Several studies suggest that self-reports of extrapyramidal symptoms yield false negative results,[41,42] but our rates are consistently similar to those reported in a multicenter study[42] of 2391 patients on neuroleptics that used self-reports and observer reports.

Dose-related adverse effects typically caused by low-potency agents—such as dry mouth—may be easily recognized and managed by a dose reduction or a switch to high-potency agents. In turn, extrapyramidal symptoms typically induced by high-potency agents—such as akathisia—imitate schizophrenic symptoms[40] and may be managed by reducing the dose but also by prescribing antiparkinsonians. That subjects on high- and combined-potency treatment received significantly higher CPZeq dosages and more psychotropic drugs would seem to support observations from at least one study[43]: when diagnosing neuroleptic-induced extrapyramidal symptoms, clinicians inevitably added an antiparkinsonian rather than reduce the neuroleptic dose.

Clinicians who prescribe neuroleptics to long-term patients may thus err on the side of overtreatment. However, the elderly as a group receive about half the neuroleptic dosage younger patients receive, which suggests clearly that clinicians recognize the need to mitigate neuroleptics' negative effects on the most fragile patients. On the other hand, clinicians appear to respond to the opposing pressure to "do everything possible" to forestall relapses.[13] Since psychotropic drugs often constitute the only tool of the prescriber, this opposing pressure might translate to a reluctance to prescribe less. □

References

1. Segal SP, Chandler S, Aviram U. Antipsychotic drugs in community-based sheltered-care homes. *Soc Sci Med.* 1980;14:589–596.
2. Hogarty GE. Depot neuroleptics: the relevance of psychosocial factors. *J Clin Psychiatry.* 1984;45:36–42.
3. Kane JM, Woerner M, Lieberman J. Tardive dyskinesia: prevalence, incidence and risk factors. In: Casey DE, Chase TN, Christensen AV, Gerlach J., eds. *Dyskinesia: Research and Treatment.* Berlin: Springer-Verlag; 1985:72–78.
4. American Psychiatric Association. APA statement on tardive dyskinesia. *Hosp Community Psychiatry.* 1985;36:902.
5. Gualtieri TC, Sprague RL. Preventing tardive dyskinesia and preventing tardive dyskinesia litigation. *Psychopharmacol Bull.* 1984;21:117–119.
6. Beck JC. Right to refuse antipsychotic medication: psychiatric assessment and legal decision-making. *Ment Phys Disability Law Reporter.* 1987;11:368–372.
7. Schussler G, Miller B. Polypharmacy in psychiatric practice. *Compr Psychiatry.* 1984;25:566–574.
8. Falloon IRH. Developing and maintaining adherence to long-term drug-taking regimen. *Schizophr Bull.* 1984;10:412–417.
9. Herz MI, Szymanski HV, Simon JC. Intermittent medication for stable schizophrenic outpatients: an alternative to maintenance medication. *Am J Psychiatry.* 1982;139:918–922.
10. Kane JM. Dosage strategies with long-acting injectable neuroleptics, including haloperidol decanoate. *J Clin Psychopharmacol.* 1986;6(suppl):205–235.
11. Kane JM. Low-dose medication strategies in the maintenance treatment of schizophrenia. *Schizophr Bull.* 1983;9:29–33.
12. Baldessarini RJ, Davis JM. What is the best maintenance dose of neuroleptics in schizophrenia? *Psychiatry Res.* 1980;3:115–122.
13. Hogarty GE, McEvoy JP, Munetz M, et al. Dose of fluphenazine, familial expressed emotion, and outcome in schizophrenia: results of a two-year controlled study. *Arch Gen Psychiatry.* 1988;45:797–805.
14. Marder SR. Depot neuroleptics: side effects and safety. *J Clin Psychopharmacol.* 1986;6:245–295.
15. Marder SR, Van Putten T, Mintz J, et al. Costs and benefits of two doses of fluphenazine. *Arch Gen Psychiatry.* 1984;41:1025–1029.
16. Kane JM. Treatment of schizophrenia. *Schizophr Bull.* 1987;13:133–156.
17. Norman TR, Judd FK, Marriott PF, Burrows GD. Pharmacotherapy of schizophrenia: clinical aspects. In: Burrows GD, Norman TR, Rubenstein G, eds. *Handbook of Studies on Schizophrenia.* Amsterdam: Elsevier; 1986;2:1–11.
18. Davis JM, Andriukaitis S. The natural course of schizophrenia and effective maintenance drug treatment. *J Clin Psychopharmacol.* 1986;6(suppl):25–105.
19. Ayd FJ. Rational pharmacotherapy: once-a-day drug dosage. *Dis Nerv Syst.* 1973;34:371–378.
20. Clark AF, Holden NL. The persistence of prescribing habits; a survey and follow-up study of prescribing to chronic hospital inpatients. *Psychiatry.* 1987;150:88–91.
21. Johnson DAW. Antipsychotic medication: clinical guidelines for maintenance therapy. *J Clin Psychiatry.* 1988;46:6–15.
22. Gelenberg AJ. Treating extrapyramidal reactions: some current issues. *J Clin Psychiatry.* 1987;48:24–27.
23. Hollister L. Strategies in clinical psychopharmacology. *Psychopharmacol Bull.* 1987;23:386–390.
24. Goldman HH, Gattozzi AA, Taube CA. Defining and counting the chronically mentally ill. *Hosp Community Psychiatry.* 1981;32:21–27.
25. Segal SP, Aviram U. *The Mentally Ill in Community-based Sheltered-Care: A Technical Supplement on Study Methodology, 1972–1977.* Berkeley, Calif: University of California; 1977.
26. Hollister L. *Clinical Uses of Psychotherapeutic Drugs.* Springfield, Ill: Thomas; 1973.
27. Schooler NR. The initiation of long-term pharmacotherapy in schizophrenia: dosage and side-effect comparison between oral and depot fluphenazine. *Pharmacopsychiatry.* 1976;9:159–169.
28. Zito JM, Craig TJ, Wanderling P, Siegel C. Pharmacoepidemiology in 136 hospitalized schizophrenic patients. *Am J Psychiatry.* 1987;144:778–782.
29. Hesse C, O'Malley K. Drug therapy in the elderly—biochemical, pharmacological, and clinical considerations. In: Bergener M, ed. *Psychogeriatrics: An International Handbook.* New York, NY: Springer; 1987:363–373.
30. American Medical Association. *AMA Drug Evaluations.* 5th ed. Chicago, Ill: AMA Division of Drugs; 1983.
31. Reardon GT, Rifkin A, Schwartz A, Meyerson A, Siris G. Changing patterns of neuroleptic dosage over a decade. *Am J Psychiatry.* 1989;146:727–729.
32. Baldessarini RJ, Katz B, Cotton P. Dissimilar dosing with high-potency and low-potency neuroleptics. *Am J Psychiatry.* 1984;141:748–752.

33. Baldessarini RJ, Cohen BM, Teicher MH. Significance of neuroleptic dose and plasma level in the pharmacological treatment of psychoses. *Arch Gen Psychiatry.* 1988;45:79–91.
34. Hayes G, Sramek J. Dosing practices for high- and low-potency neuroleptics. *Am J Psychiatry.* 1988;145:769.
35. Van Putten T, Marder SR, Mintz J, Poland RE. Haloperidol plasma levels and clinical response: a therapeutic window relationship. *Psychopharmacol Bull.* 1988;24:172–175.
36. Quitkin F, Rifkin A, Klein DF. Very high dosage vs. standard dosage fluphenazine in schizophrenia: double-blind study of non-chronic treatment-refractory patients. *Arch Gen Psychiatry.* 1975;32:1276–1281.
37. Donlon PT, Hopkin J, Tupin JP. Overview: efficacy and safety of the rapid neuroleptization method with injectable haloperidol. *Am J Psychiatry.* 1979;136:273–278.
38. Van Putten T, Marder SR. Behavioral toxicity of antipsychotic drugs. *J Clin Psychiatry.* 1987;48:13–19.
39. Kane JM, Rifkin A, Woerner M, et al. Low dose neuroleptic treatment of outpatient schizophrenics: I. Preliminary results for relapse roles. *Arch Gen Psychiatry.* 1983;40:893–896.
40. Marder SR, Van Putten T, Mintz J, Lebelle M, McKenzie J, May PRA. Low- and conventional-dose maintenance therapy with fluphenazine decanoate: two year outcome. *Arch Gen Psychiatry.* 1987;44:518–521.
41. Alexopoulos GS. Lack of complaints in schizophrenics with tardive dyskinesia. *J Nerv Ment Dis.* 1979;167:125.
42. Dencker SJ, Ahlfors UG, Bech P, Elgen K, Lingjaerde O. Classification of side effects in psychopharmacology. *Pharmacopsychiatry.* 1986;19:40–42.
43. Weiden PJ, Mason JJ, Haas G, Mattson J, Frances A. Clinical nonrecognition of neuroleptic-induced movement disorders: a cautionary study. *Am J Psychiatry.* 1987; 144:1148–1153.

Eligibility for Publicly Financed Home Care

ABSTRACT

Objectives. Proposals for publicly financed home care for the elderly now tend to include cognitive impairment criteria as well as activities of daily living (ADL) criteria. The numbers of elderly deemed eligible for services will depend on the definitions of ADL and cognitive impairment used.

Methods. Data from the 1984 National Long-Term Care Survey were used to generate a series of estimates of the community-dwelling elderly with ADL disabilities and cognitive impairment.

Results. When only ADL criteria are used, estimates of disability range from 472 000 to over 3 million (1.6% to 12.5% of the community-dwelling elderly). These estimates increase to approximately 1 million to 4.2 million (3.5% to 14.0% of the community-dwelling elderly) when cognitive impairment criteria are added.

Conclusions. The use of more stringent or more liberal eligibility criteria will have dramatic effects on the number of elders who qualify for services. The nature of the eligibility criteria employed in any expansion of federally financed home care benefits will be a major factor in determining the costs of such a program. (*Am J Public Health.* 1992;82:853–856)

Mary E. Jackson, PhD, Brian Burwell, BA, Robert F. Clark, DPA, and Mary Harahan, MA

Introduction

At present, there is no uniform federal policy for financing home care services for elderly persons in need of long-term care. Eligibility criteria that determine who is eligible to receive publicly financed home care services varies considerably by funding source. To a large extent, these criteria are established at the state level, because states have primary responsibility for allocation decisions under many of these funding mechanisms, particularly Medicaid. However, recent congressional proposals to expand federal coverage of home care have stipulated with increasing specificity uniform criteria for eligibility based on functional disability. Typically, these criteria rely on measures of disabilities in a given number of activities of daily living (ADLs).

Eligibility rules that rely exclusively on functional disability criteria have been criticized for excluding persons who have the physical capacity to perform ADLs, but not the ability to remember when to initiate a task, and those who need supervision or cuing in order to complete a task. These are typically individuals who have Alzheimer's disease or a related disorder and who are cognitively impaired. Most legislative proposals for expanding federal financing for home care now include criteria intended to capture the cognitively impaired population targeted for services.

Although nearly all proposals have been specific about the number of ADL disabilities necessary to qualify a person as eligible for services, only a few go so far as to define how disability in a specific ADL would be measured. If implemented, any of these proposals would have to develop more precise criteria for defining an ADL disability in order to ensure equitable access to program applicants. As Stone and Murtaugh have demonstrated, estimates for the number of persons with disability can range widely depending upon the ADL definition invoked.[1]

As nonspecific as some of the ADL eligibility criteria in the expansion proposals are, they are vastly more specific than language pertaining to criteria that would establish eligibility based on cognitive impairment. This lack of specificity regarding cognitive criteria is not simply a deficiency in policy development, but reflects the state of the art in measuring and screening for cognitive impairment in the elderly population.

To demonstrate how cognitive impairment may be operationalized within the context of eligibility criteria, we present estimates of the number of functionally disabled *and* cognitively impaired community-dwelling elders. We vary both the definition of functional (ADL) disability and the definition of cognitive impairment, and show how sensitive estimates can be to the criteria on which they are based.

Mary E. Jackson and Brian Burwell are with SysteMetrics, Lexington, Mass. Robert F. Clark and Mary Harahan are with the US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, Washington, DC.

Requests for reprints should be sent to Mary E. Jackson, PhD, SysteMetrics, 24 Hartwell Avenue, Lexington, MA 02173.

This paper was submitted to the Journal January 18, 1991, and accepted with revisions November 12, 1991.

The views and conclusions presented in this paper are solely those of the authors and do not necessarily reflect the views of the Department of Health and Human Services.

Methods

Data

Data from the 1984 National Long-Term Care Survey (NLTCS) were used to generate estimates of functional disability and cognitive impairment. As part of the NLTCS, in-depth interviews were conducted of a sample of Medicare beneficiaries aged 65 years and older on April 1, 1984, who were found to have either an ADL or an Instrumental Activity of Daily Living (IADL) limitation.[2]

Estimates are based on the community-resident (noninstitutionalized) sample of the NLTCS. They have been subjected to a weighting procedure whereby cross-sectional weights were applied to adjust for nonresponse and the complexities of the sampling design. Application of weights results in nationally representative estimates of the community-dwelling population aged 65 years and older. All estimates presented in this paper meet precision standards of relative standard errors of 30% or less, and are based on unweighted cell sizes of 400 or larger. Because the estimates derived from the survey reflect the number of persons disabled in 1984, we have adjusted 1984 estimates to be consistent with 1990 population estimates for the population aged 65 years and older.

ADL Disability Measures

Estimates of functional disability are based on the same alternative definitions of ADL disability as used by Stone and Murtaugh, and the methods used to derive them are replications of the methods of Stone and Murtaugh.[1] The estimates were developed from specific questions posed to NLTCS respondents or their proxies regarding the types of help the respondent received in five core ADLs: bathing, dressing, toileting, transferring, and eating. Receipt of services is used in these analyses as an indicator of disability.

The NLTCS asked multiple questions about functioning in each of the five ADLs. Disability in ADL was defined in three different ways, according to the type of assistance required to perform a given ADL.

- Receives active assistance to perform an ADL
- Receives active or standby assistance to perform an ADL
- Receives active or standby assistance to perform an ADL or uses an assistive device in performing an ADL

All estimates refer to ADL disabilities that were chronic in nature, that is, those that had lasted for 3 months or longer.[1]

Cognitive Impairment

The NLTCS was not specifically designed to estimate the number of elders with cognitive problems; thus we were able to derive cognitive impairment estimates only for persons who met the functional disability screen for the NLTCS sampling frame. However, the NLTCS sampling frame for identifying disabled community elders was quite liberal in its definition of disability. Any type of assistance, human or mechanical, for any ADL (i.e., core ADLs plus bladder or bowel incontinence, indoor mobility, outdoor mobility) or for any IADL (meal preparation, laundry, light housework, grocery shopping, money management, medication management, telephoning) that had lasted, or was expected to last, for a minimum of 90 days was included in the definition. Given the wide net that the NLTCS cast in its sampling frame of the elderly disabled, as well as the fact that most cognitively impaired elders can expect to have at least some limitation in functioning, we are fairly confident that elders with cognitive impairments are sufficiently represented in the estimates that follow.

Three definitions of cognitive impairment were developed. The first includes persons with an error score of 4 or more on a modified Short Portable Mental Status Questionnaire (SPMSQ). The SPMSQ is a 10-item screening instrument designed to identify elderly persons with probable cognitive impairment.[3] A score of 4 or higher, rather than the customary 5 or higher, was used to indicate cognitive impairment because one of the SPMSQ questions was altered in the NLTCS, resulting in fewer errors than the original item typically elicits. The cutoff point indicating cognitive impairment was adjusted downward to accommodate this modification. This measure of cognitive impairment is referred to as "CI" in Table 1.

SPMSQ data are missing for those sample members who were not capable of responding to the questionnaire and whose information was provided by a proxy. To identify sample members with missing SPMSQ information who were cognitively impaired, alternative criteria were developed. If a person with missing SPMSQ information was judged to be senile by the respondent's proxy, or if the proxy reported that the person had a disability caused by disorders classified as mental retardation or senile psychotic conditions, the person was considered to be cognitively impaired.[4]

The second definition of cognitive impairment, an attempt to operationalize eligibility language from some legislative proposals, includes evidence of behavior problems along with the CI criteria described above. Some argue that behavior problems should be considered in determining the need for long-term care services because such problems require supervision and oversight to ensure the health and safety of the afflicted person.

Wandering, frequent temper tantrums, and compulsive stealing are the behaviors on which the NLTCS collected information, and are therefore the behaviors represented in our behavior problems (BP) measure. Because the NLTCS did not collect information on other types of behavior problems, our results may underestimate the prevalence of behavior problems among the elderly.

The third definition of cognitive impairment employed in our estimates is more stringent than the other two. Under the "CI+" measure, a person must not only meet the CI criteria but must also either demonstrate disability (i.e., must receive active or standby assistance for 3 months or longer) in at least one ADL or one of the cognitively oriented IADLs (money management, medication management, telephoning) *or* exhibit a behavior problem. For a person to be considered cognitively impaired under this definition, there must be evidence of the need for oversight or some type of assistance. The CI+ criteria is an attempt to operationalize language in proposed legislation that stipulates that a person with Alzheimer's disease or other cognitive impairments must demonstrate needs similar to those of persons meeting ADL criteria in order to be deemed eligible. The CI+ criteria operationalize such criteria by requiring some evidence of the need for care beyond a diagnosis or a score on a cognitive-impairment screen. A diagnosis or a score may not reflect the fact that an impairment is severe enough to require home care services.

Estimates

Table 1 presents estimates of the number of disabled community-dwelling elderly based on alternative ADL as well as alternative cognitive impairment criteria. In each section, the first column represents the number of persons estimated to be disabled based on functional criteria alone. The remaining columns reference

estimates of the number of ADL-disabled *or* cognitively impaired persons; the differences in estimates between the last three columns are a function of the definition of cognitive impairment invoked.

For example, the estimate in the first column of section A for 2+ ADLs represents the number of elders estimated to have two or more ADL disabilities, where disability is defined as the need for the active assistance of another person to perform the ADLs in question.

The estimate in the second column references the number of elders who would be eligible for services if the criterion were 2+ ADLs *or* CI *or* BP. This decision rule targets individuals with two or more ADL disabilities (active assistance only) *or* those who are cognitively impaired (those who score 4+ on the SPMSQ or who have a cognitive impairment–related diagnosis) *or* those who exhibit one or more behavior problems (wandering, frequent temper tantrums, or compulsive stealing). The rationale for considering those with behavior problems within the cognitive impairment group is that behavior problems are considered an indicator of cognitive impairment, and those with behavior problems are thought to need oversight and supervision to ensure their health and safety.

Slightly more than 2 million elders would meet the 2+ ADLs (active only) *or* CI *or* BP criterion in 1990. Approximately 1 million additional persons (or 3.5% more of the elderly population) would be eligible for services if this criterion were used instead of comparable criteria that exclude the cognitively impaired from consideration.

The "CI or BP" criterion is the most liberal of the three cognitive impairment criteria included in the estimates; it clearly yields the highest estimates of all the measures. However, this particular criterion may result in an overestimate of the cognitively impaired population because of the way in which the behavior problem factor figures into the criterion. For example, under the 2+ ADLs *or* CI *or* BP rule a person without any ADL disabilities, who is not considered cognitively impaired according to the CI criterion, but who compulsively steals, would be deemed eligible. A person with this particular profile may not be an appropriate candidate for home care services, and including such persons in an eligibility pool may result in misspecifying the targeted population, even under very liberal targeting policies.

A more conservative approach would be to define cognitive impairment solely in terms of evidence of cognitive impairment, as exemplified by the CI criterion. With this criterion, fewer individuals would fall into the eligibility pool. Although using the cognitive impairment criterion instead of the CI *or* BP criterion does lower the estimates, the decrease is only on the order of 0.2% to 0.5%, or 60 000 to 150 000 fewer eligible persons nationwide, depending upon the exact ADL criteria invoked and the definition of disability used.

This approach, though more conservative, is still subject to criticism. A score on a cognitive impairment test, or even a diagnosis of Alzheimer's disease, does not necessarily correspond to a need for home care. The target population for home care services should probably not include persons for whom organicity has not yet impeded independent functioning of ADLs or presented a threat to the safety of the individual. This criticism has led to the suggestion that eligibility for home care services for cognitively impaired persons be linked to more specific evidence of the need for care. The CI+ criterion responds to this suggestion by requiring that eligibility for services be evidenced by an SPMSQ score indicating cognitive impairment (or a cognitive impairment–related diagnosis) *and* some evidence of the need for care or oversight. Such care or oversight might be required because of (1) one or more ADL disabilities *or* (2) a disability in one of the three cognitively oriented IADLs (money management, medication management, or telephoning) *or* (3) display of behavior problems.

The CI+ criterion is undoubtedly the most restrictive of the three cognitive impairment criteria. For example, 6.8% of the nation's elderly would be service-eligible under the 2+ ADLs (active) *or* CI *or* BP criterion, and 6.5% under the 2+ ADLs *or* cognitive impairment criteria, but only 4.8% under the 2+ ADLs *or* CI+ criterion. Yet this 4.8% estimate is substantially higher than would result from the same criterion without the cognitive impairment component added. This translates into an additional 450 000 eligible persons under the 2+ ADLs *or* CI+ criterion who are eligible because of the cognitive impairment component alone.

TABLE 1—Number of Community Elderly (in Thousands) Aged 65+, by Alternative Activities of Daily Living (ADL) Disability and Cognitive Impairment Criteria, 1990

No. ADLs[d]	No. Meeting ADL Criteria	No. Meeting ADL or Cognitive Impairment Criteria: CI or BP[a]	CI[b]	CI+[c]
A. ADL Disability Criterion: Active Assistance Only				
1+	1837	2701	2606	2114
	(6.1)	(9.0)	(8.7)	(7.0)
2+	1007	2056	1947	1458
	(3.3)	(6.8)	(6.5)	(4.8)
3+	630	1780	1657	1174
	(2.1)	(5.9)	(5.5)	(3.9)
B. ADL Disability Criterion: Active or Standby Assistance				
1+	2417	3179	3098	2619
	(8.0)	(10.6)	(10.3)	(8.7)
2+	1578	2250	2425	1942
	(5.2)	(8.4)	(8.1)	(6.4)
3+	1157	2182	2076	1588
	(3.8)	(7.3)	(6.9)	(5.3)
C. ADL Disability Criterion: Active or Standby Assistance or Assistive Device				
1+	3770	4217	4164	3915
	(12.5)	(14.0)	(13.8)	(13.0)
2+	2455	3192	3107	2717
	(8.2)	(10.6)	(10.3)	(9.0)
3+	1561	2490	2390	1946
	(5.2)	(8.3)	(7.9)	(6.5)

Note. Numbers in parentheses represent the percentage of the community-residing population aged 65 years and older who meet the criteria.

[a]Score of 5+ on Short Portable Mental Status Questionnaire (SPMSQ) or cognitive impairment–related diagnosis or behavior problem (BP).

[b]Score of 5+ on SPMSQ or cognitive impairment–related diagnosis.

[c]CI plus behavior problems or CI plus either 1+ ADL or disability in money management, telephoning, or medication management.

[d]There are five core ADLs: bathing, dressing, toileting, transferring, eating. The numbers in this column indicate how many of these activities with which the individual needs help.

Discussion

In this study, we have demonstrated how various eligibility criteria can be used to derive estimates of the disabled elderly population living in the community, taking into account those in need of services because of cognitive impairment. We have shown how these various criteria affect the size of the population that would be made eligible for benefits under alternative definitions of functional disability and cognitive impairment.

It must be recognized that these estimates are derived from survey data, and incentives for survey respondents to overstate their disabilities are minimal. Survey respondents, unlike program applicants, receive no benefit from exaggerating their disability. In fact, there is some evidence that survey respondents may underestimate their disability levels.[5] Thus we would expect some increase in self-reported disability if persons were applying for services and knew that their receipt of service was contingent upon a disability measure. Additionally, when service eligibility is predicated on the need for *human* assistance in performing ADLs, there may be a disincentive to substitute assistive devices for human help, thus rendering the impaired more dependent than necessary and also jeopardizing the development of new technologies for fostering independence.

Nonetheless, the estimates presented in this paper demonstrate how substantially different estimates of the disabled elderly population with and without cognitive impairment, can be generated by employing varying definitions of disability. The use of more stringent or more liberal eligibility criteria for home care services will have dramatic effects on the number of elders who would qualify for services. The most obvious implication is that the specific eligibility criteria employed in legislation to expand federally financed home care services will be major factors in determining the ultimate costs of such an expansion. A second implication is that unless the eligibility criteria designated in the legislation are specific in their definitions, the number of persons who will actually qualify for benefits will depend largely upon how the general language of the legislation is interpreted in regulatory and administrative procedures during program implementation. This could lead to considerable discrepancy between the intent of the legislation and its actual implementation.

Definitions of disability using measures of ADL impairment have become popular mechanisms for allocating home care resources under proposed federal legislation. However, the need for publicly financed home care depends upon a variety of factors, not just one's level of disability. Financial need is an obvious factor, and one that is also built into many legislative initiatives. However, it is also true that the elderly adjust to their disabilities through a multitude of adaptive behaviors, just as nonelderly persons with disabilities do. In addition, many elderly persons have access to support from family and friends that can augment the formal provision of services. Given the variety of factors that contribute to the need for home care, should a federal entitlement program for home care services, based solely on standardized measures of disability, be implemented? What role, if any, should the availability of family and friends, as well as financial status, play in eligibility determination and service allocation?

In any case, it is apparent that the establishment of criteria to determine which elderly persons with disabilities should be eligible for publicly funded home care and which should not is not a simple matter. In this context the question must be raised whether it is possible to enact federal legislation that would ensure that benefits are allocated in an equitable manner according to true need. Unlike access to publicly financed acute care services, which is determined primarily by physicians, access to expanded federal home care benefits will require policy makers to establish relatively specific decision rules about who is eligible for benefits and who is not.

Acknowledgments

Funding for this study was provided under contract HHS-100-88-0041 from the Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, Division of Disability, Aging, and Long-Term Care.

We would like to acknowledge the assistance of William D. Spector, PhD, and Peter Kemper, PhD, from the Agency for Health Care Policy and Research for providing some of the analytic files used in these analyses. We would also like to thank Sidney Katz, MD, for his review of an earlier draft of this manuscript. The technical assistance provided by Karen Koch of Social and Scientific Systems and Sean Kennedy of SysteMetrics is also greatly appreciated.

References

1. Stone R, Murtaugh C. The elderly population with chronic functional disability: implication for home care eligibility. *Gerontologist.* 1990;30:491–496.
2. Manton KG. A longitudinal study of functional change and mortality in the United States. *J Gerontology.* 1988;43:S153–161.
3. Pfeiffer E. A Short Portable Mental Status Questionnaire for the assessment of organic brain deficit in elderly patients. *J American Geriatrics Society.* 1975;23:433–441.
4. Spector W. Cognitive impairment and disruptive behavior among elderly persons living in the community: implications for targeting long-term care. *Gerontologist.* 1991; 31:51–59.
5. Ford S, Folmar SJ, Salmon RB, Medalie JH, Roy AW, Galazka SS. Health and function in the old and very old. *J American Geriatrics Society.* 1988;36:187–197.

Postmarketing Surveillance Study of Permethrin Creme Rinse

Elizabeth B. Andrews, PhD, Michael C. Joseph, MD, MPH, Mark J. Magenheim, MD, MPH, Hugh H. Tilson, MD, DrPH, Peggy A. Doi, and Margaret W. Schultz, MPH

ABSTRACT

Background. An observational, epidemiological study was undertaken to evaluate the safety of permethrin 1% creme rinse (Nix™) for treatment of head lice infestations.

Methods. Thirty-seven local public health departments enrolled a total of 38 160 patients for 47 578 treatments with permethrin or other pediculicides from September 1, 1986, through January 31, 1988. Follow-up safety information was collected between 7 and 14 days following treatment via return visit or telephone contact.

Results. One hundred three adverse events were reported among 41 955 evaluable treatments. The rates of reported adverse events were 2.2 per 1000 treatments among permethrin treatments, 3.4 per 1000 treatments among lindane treatments, and 1.5 per 1000 treatments among other over-the-counter treatments. No serious, unexpected adverse events were detected in the 18 950 patients treated with permethrin.

Conclusions. This study confirmed the safety profile of permethrin in conditions of general use, as seen in clinical trials. Postmarketing safety monitoring in public health departments of drugs used to treat public health conditions was shown to be feasible. (*Am J Public Health.* 1992;82:857–861)

Introduction

Infestation with head lice continues to be a public health concern in the United States, particularly among schoolchildren. Surveys indicate that approximately 2.6 million households (or 2.8 million individuals) experience an outbreak of head lice each year, and studies by the Centers for Disease Control between 1973 and 1974 found an overall prevalence of head lice of approximately 8% in schoolchildren (grades K to 8).[1–3]

Permethrin 1% creme rinse, hereafter referred to as permethrin, contains the synthetic pyrethrin permethrin (a pyrethroid). It has been available by prescription for head lice treatments since June 1986. Before its marketing, the only available pediculicides were prescription compounds containing lindane (Kwell and generic products) and over-the-counter (OTC) products, most of which contain pyrethrins (A–200®, RID®, and others). Since completion of this study, permethrin has become available OTC.

In clinical studies of over 3500 patients, the only expected adverse drug experience associated with permethrin was local irritative symptoms, which ranged in frequency from 2.1% to 5.9%. Many of these symptoms were thought to be the result of the lice infestation.[4–6]

This postmarketing safety study provided the principal new data on which the conversion of permethrin from prescription to OTC status was based.[7] It used recognized epidemiological methods both to assess the short-term safety of permethrin for general use in the treatment of head lice and to identify, document, and assess rare but medically important events following treatment. It was not intended to compare effectiveness of the different pediculicides.

Methods

In this observational, epidemiological safety study, public health departments dispensed permethrin or another pediculicide, counseled about treatment in the customary methods, and followed each patient between 7 and 14 days after treatment via return visit or telephone call to identify any subsequent adverse clinical events.

Health departments developed treatment allocation schemes to ensure that each diagnosed individual had essentially the same chance of being treated with permethrin or another pediculicide. This quasi-randomization system included such measures as assigning treatment according to day of the week patient was diagnosed with head lice infestation (there being no a priori reason to believe that patient mix varied by day of week), and alternating treatment choice by patient or family group. Analyses of baseline characteristics, by site, were conducted to evaluate the adequacy of the allocation schemes.

The authors are with the Burroughs Wellcome Company at Research Triangle Park in North Carolina and the Sarasota County Health Department in Florida.

Requests for reprints should be sent to Elizabeth B. Andrews, PhD, Burroughs Wellcome Company, 3030 Cornwallis Road, Research Triangle Park, NC 27709.

This paper was submitted to the Journal May 11, 1990, and accepted with revisions July 10, 1991.

Health departments enrolled infested school-aged children and their family members ages 2 years or above with head lice infestation. Pregnant women and persons with a history of allergic reactions to pyrethrins were excluded.

The existing staff of 37 public health departments in 12 states were collaborative study investigators (see acknowledgement). During the treatment period (September 1, 1986, to January 31, 1988), 38 160 patients were enrolled for 47 578 treatments. Four sites in Michigan enrolled 13% of all patients, 10 other sites in 10 states enrolled 27%, and 23 sites in Florida enrolled the remaining 60%. The median enrollment for individual sites was approximately 650 patients (ranging from 65 to 3293).

In July 1986, the State of Florida Institutional Review Board (IRB) of the Department of Health and Rehabilitative Services and an independent, federally qualified IRB of the fourth-largest hospital in Florida (an 862-bed public facility with substantial research activities) reviewed and approved the full research protocol, including methods and detailed provisions for the protection of human subjects, for all participating Florida study sites. All participating health departments were provided permethrin free of charge and were reimbursed for study expenses, including the cost of obtaining other pediculicides, according to negotiated budgets. Further details of the process of designing and implementing a network of collaborating public health clinic sites will be described in another paper.

Treatment and Assessment Plan

At enrollment, patient initials, birth date, sex, ethnic group, history of allergic problems, name of the pediculicide dispensed, and previous enrollment in the study were recorded.

Follow-up information was solicited using a common protocol. If patients did not return to the health department, a telephone callback system was used to gather the information. When an adverse event was reported, symptoms, concurrent medications and illnesses, and course of required treatment, if any, were solicited. An adverse event was classified as "medically important" if medical consultation was actually sought or if the event was likely to require clinical monitoring, whether or not any medical consultation was obtained. Events of unspecified or ambiguous severity were classified as medically important.

Information was obtained on treatments for 38 180 patients. Only those treatments for which follow-up was complete (41 955 treatments of 34 275 patients) were used in the analysis. The proportion of patients available for follow-up did not vary by treatment type, and their baseline characteristics did not differ from those of patients lacking follow-up.

Classified by pediculicide exposure, 18 950 (55.3%) patients received permethrin, 10 155 (29.6%) received lindane products, and 5170 (15.1%) received one of the 28 different OTC products. OTC products included RID® (3171 patients), Pyriny!® (699 patients), and an assortment of similar pediculicides. Lindane and OTC groups were analyzed separately because of differences in their potency and safety profiles.

Sample Size and Power

The sample size was sufficient to detect a threefold increase in the incidence of an adverse event that had an expected frequency of 1 in 1000, with a 95% significance level when comparing the group treated with permethrin with the group treated with lindane, and 80% power and a 95% significance level when comparing the permethrin group with the group treated with OTC products. Calculations were based on the binomial distribution.

Results

Participant Characteristics

The distribution of patients by age, sex, ethnic group, and allergy history was similar among the three treatment groups (Table 1). No significant differences were seen across sites (data available upon request). The age distribution of the study participants was bimodal: 57% were between 2 and 15 years of age and 25% were between 21 and 55 years of age. Eight percent of the participants reported a history of allergies or allergic conditions.

Analysis Using All Treatment

Data were first analyzed using all treatments for which follow-up information was obtained (n = 41 955). Because the retreatment experience might be dependent on the outcome of the first study treatment, an analysis of data restricted to the first study treatment was also completed (n = 34 275).

Of the 41 955 evaluable treatments, there were 103 reported adverse events (2.5 per 1000), 2.2 per 1000 (95% CI: 1.6, 2.9 per 1000) after permethrin treatment, 3.4 per 1000 (95% CI: 2.5, 4.6 per 1000) after lindane treatment, and 1.5 per 1000 (95% CI: 0.7, 2.7 per 1000) after OTC treatment. The risk difference when comparing permethrin with lindane was –1.2 per 1000 (i.e., 1.2 fewer problems per 1000 permethrin treatments; 95% CI: –2.4, –0.05). The risk difference when comparing permethrin with OTC treatments was 0.7 problems per 1000 treatments (i.e., 0.7 more problems per 1000 permethrin treatments; 95% CI: –0.4, 1.8 per 1000).

Consistent with results from clinical trials in the United States, all three treatment groups reported pruritus and rash more frequently than they reported other adverse events. Eight events were considered medically important. When the statistical analysis was restricted to those events, or to events for which medical attention was sought, there were no significant differences between the permethrin group and either of the comparison groups. In the permethrin group, five adverse events, summarized below, were medically important. Two of the patients were not treated with permethrin themselves but came into contact with a treated family member.

Case P1: A 16-year-old girl with no previous history of allergy experienced breathing difficulty while her sister was being treated with permethrin. The episode subsided after the girl left the treatment room.

Case P2: A 34-year-old woman with a history of drug allergies experienced mild nausea and moderate to severe shortness of breath that lasted about 4 hours. The patient had been taking Inderal® (propranolol) (12 mg daily).

Case P3: A 10-year-old girl with a history of asthma and food allergies experienced arm shaking and stuttering that lasted less than a minute and did not recur. The patient did not require a medical visit. She had used the pediculicide A-200® on two occasions, 10 days and 6 days before treatment with permethrin.

Case P4: A 35-year-old woman (mother of P3) with a history of food allergies experienced two incidents of arm shaking for about 40 seconds and some skin irritation. The patient did not have a medical visit for these events. The Health Department physician stated that the arm shaking was probably due to the patient's steady use of her arms in applying 10-minute massages to three individuals on the same day.

TABLE 1—Characteristics of Study Population

	Permethrin[a] (n = 18 950)		Lindane (n = 10 155)		OTC[a] (n = 5170)		Total (n = 34 275)	
	No.	(%)	No.	(%)	No.	(%)	No.	(%)
Mean age (y)	18.5		18.4		16.9		18.3	
Sex[b]								
Male	7938	(42.0)	4332	(42.8)	2122	(41.1)	14 392	(42.1)
Female	10 982	(58.0)	5795	(57.2)	3044	(58.9)	19 821	(57.9)
Ethnic group[c]								
Caucasian	17 161	(90.8)	9094	(89.6)	4527	(87.9)	30 782	(90.1)
Hispanic	1140	(6.0)	762	(7.5)	418	(8.1)	2320	(6.8)
Native American	286	(1.5)	82	(0.8)	99	(1.9)	467	(1.4)
Black	95	(0.5)	116	(1.2)	37	(0.7)	248	(0.7)
Asian	55	(0.3)	49	(0.5)	4	(0.1)	108	(0.3)
Other	162	(0.9)	25	(0.3)	66	(1.3)	253	(0.7)
Positive allergy history	1602	(8.5)	790	(7.8)	351	(6.8)	2743	(8.0)
Drug allergy	624	(3.3)	334	(3.3)	112	(2.2)	1070	(3.1)
Asthma	347	(1.8)	130	(1.3)	110	(2.1)	587	(1.7)
Hay fever	208	(1.1)	98	(1.0)	40	(0.8)	346	(1.0)
Food allergy	112	(0.6)	44	(0.4)	14	(0.3)	170	(0.5)
Eczema	61	(0.3)	32	(0.3)	18	(0.3)	111	(0.3)
Hives	31	(0.2)	19	(0.2)	5	(0.1)	55	(0.2)
Other	219	(1.2)	133	(1.3)	52	(1.0)	404	(1.2)
Adverse events								
Initial treatments only	46	(0.2)	36	(0.4)	5	(0.1)	87	(0.3)
All treatments	49	(0.2)	44	(0.3)	10	(0.1)	103	(0.3)
Total treatments	22 362		12 856		6737		41 955	

[a]OTC = over-the-counter products.
[b]Data on sex missing for 82 patients.
[c]Data on ethnic group missing for 97 patients.

Case P5: A 23-year-old woman with no previous allergy history experienced a swollen face beginning 1 day after treatment and lasting 1 week. She required medical treatment. The patient reported that the adverse event could have been caused by a new facial soap and that she did not attribute it to the permethrin treatment.

Among patients treated with lindane, three adverse events were judged medically important.

Case L1: A 31-year-old woman with no previous history of allergies experienced "extreme" blurred vision that lasted 7 days and did not recur. The event required a medical visit.
Case L2: A 39-year-old man with an extensive history of allergies and asthma experienced moderate respiratory irritation that lasted 18 hours. No medical visit was required.
Case L3: A 28-year-old woman with a history of asthma, hay fever, and other allergies experienced moderate difficulty in breathing for 24 hours following treatment. An OTC bronchodilator was used to treat the episode, and no medical visit was required.

Among patients treated with OTC pediculicides, there were not medically important adverse events.

Analysis Using First Treatment per Patient

Among the 34 275 first study treatments, there were 87 reported adverse events, 2.4 per 1000 (95% CI: 1.8, 3.2 per 1000 patients) after permethrin treatment, 3.5 per 1000 (95% CI: 2.5, 4.9 per 1000 patients) after lindane treatment, and 1.0 per 1000 (95% CI: 0.3, 2.3 per 1000 patients) after OTC treatment. The risk difference when comparing permethrin with lindane was –1.1 per 1000 (i.e., 1.1 fewer problems per 1000 patients treated with permethrin; 95% CI: –2.5, 0.2 per 1000). The risk difference when comparing permethrin with OTC patients was 1.5 events per 1000 patients (i.e., 1.5 more problems per 1000 patients treated with permethrin; 95% CI: 0.4, 2.6 per 1000). When the analyses were restricted to adverse events that appeared to be medically important or for which medical consultation was sought, there were no statistically significant differences between the permethrin group and either the lindane or the OTC group.

Stratified analyses were performed to determine if the different reporting rates for adverse events were caused by small differences between treatment groups in age, sex, ethnicity, or allergy status. The prior findings were not altered substantially when these factors were controlled (data available upon request).

Analyses were also performed across all exposure groups to assess the relationship between patient characteristics and the reporting of adverse events following treatment. When all pediculicidal exposure was combined (Table 2), patients with a history of allergy were more than four times as likely (8.4 per 1000) to report an adverse event after treatment than were patients without a prior allergy history (2.0 per 1000). The risk difference was 6.4 per 1000 (95% CI: 2.9, 9.8 per 1000). In addition, females were 2.4 times as likely (3.4 per 1000) to report an adverse event than males (1.4 per 1000). The risk difference was small (2.0 per 1000 patients; 95% CI: 0.98, 3.00 per 1000).

Discussion

The overall rate of adverse events reported after treatment in this study was very low—between 1.0 and 3.5 events per 1000 patients. This rate of events was approximately one tenth that seen in more

TABLE 2—Adverse Events by Allergy History and Sex: All Exposure Groups Combined Using First Treatment per Patient

	Total No. Events	No. Patients Reporting Adverse Events	% Reporting Adverse Events	(95% CI)
Allergy history*				
Any allergy	2743	23	0.84	(0.53, 1.26)
No allergy	31 447	63	0.20	(0.15, 0.26)
Specific allergy				
Drug allergy	1070	4	0.37	(0.10, 0.95)
Asthma	587	5	0.86	(0.28, 1.98)
Hay fever	346	4	1.16	(0.32, 2.93)
Food allergy	170	3	1.77	(0.32, 5.07)
Eczema	111	2	1.80	(0.22, 6.36)
Hives	55	1	1.82	(0.05, 9.72)
Other	404	4	0.99	(0.27, 2.52)
Sex**				
Male	14 392	20	0.14	(0.09, 0.22)
Female	19 821	67	0.34	(0.26, 0.43)

*Allergy history not reported for 85 patients.
**Sex not reported for 62 patients.

closely monitored experimental trials in the United States. Because this postmarketing epidemiological study sought primarily to discover rare, serious, and unexpected adverse events, data collection was not designed to assess or confirm frequent and less serious events. Therefore, the low reporting rate of common, expected events observed in this study was predictable.

Although adverse events might have been underreported, there was no incentive for participants to report preferentially or underreport problems by treatment group. However, it is acknowledged that such bias could not have been strictly controlled in this unblinded, nonrandomized trial. The perceived risk of treatment with each product by patients, parents of patients, and medical staff, or the perception that this was a "permethrin study," could have introduced bias in reporting problems.

When the rate of overall reported problems among the permethrin treatments was compared with the rates among the lindane and OTC groups, there was a consistent pattern. The permethrin patients reported a rate of adverse events lower than that of the lindane patients and higher than that of the OTC patients. Observed differences, however, were small. Because the sample size was so large, even small and perhaps unimportant differences would be statistically significant. Moreover, risk differences observed in this study were small, on the order of one per thousand patients or treatments.

There were no significant differences between permethrin and either the lindane or the OTC comparison group when the analysis was restricted to those problems that were medically important. The most important adverse event was difficulty in breathing, which was reported in 4 patients (2 of 18 950 permethrin patients and 2 of 10 155 lindane patients), 3 of whom had a previous history of allergies.

No serious or unexpected adverse event was detected in association with permethrin despite a sample size adequate to detect such an event at a rate of 1 per 10 000 treatments with 85% probability.

Although considerations for patient safety and methodology were paramount in the study design and analysis, attention to feasibility and to patient and provider acceptability were also important practical objectives. Careful attention not to interfere unduly with normal patient care activities imposed several limitations on study design (and on data quality to a minimal degree). Strict randomization of pediculicide treatments was not feasible, and the choice of allocation schemes varied among and within participating local public health agencies. However, once determined, individual allocation schemes were reported to have been followed, and there was no evidence of systematic bias in overall treatment allocation, given that summary analyses indicated comparable allocations among all treatment options. Additional analyses comparing patient characteristics and results by site also showed no differences.

Although complete follow-up was not attainable for 10% of enrolled patients, there is no a priori reason to believe that the patients lacking follow-up information experienced different outcomes from those experienced by patients who provided such information. Analyses of baseline characteristics showed that the two groups were not different. Furthermore, because the local health department is often an important source of medical care in the community, it is likely that significant problems among patients without explicit follow-up information would have come to the attention of the responsible health department.

This public-private collaborative study approach demonstrates the ability to gather relatively basic, yet clinically important safety information on an extremely large and diverse study population through a public health field research stratagem without unduly compromising safety or data quality. □

Acknowledgments

This study was funded by research grants from the Burroughs Wellcome Company. The authors wish to thank the participating health departments as well as Kathy Schoeffler, Linda Wastila, and Anita Blanchard for their assistance in data management and quality assurance. The participating health departments: La Paz County, Arizona; Dougherty County, California; San Bernardino County, California; these Florida counties: Alachua, Collier, Gadsden, Hardee, Highlands, Hillsborough, Jackson, Lake, Lee, Pasco, Pinellas, Polk, Sarasota, Sumter, Putnam, Calhoun, DeSoto, Escambia, Brevard, Nassua, St. Johns, Okaloosa, Holmes; Winnebago County, Illinois; Marion County, Indiana; Branch-Hillsdale-St. Joseph District, Michigan; Midland, Michigan; Muskegon County, Michigan; Ottawa County, Michigan; CUHCC, Minnesota; Moore County, North Carolina; Nashua, New Hampshire; Warren City, Ohio; Bremerton-Kitsap County, Washington.

References

1. Slonka GF, McKinley TW, McCroan JE, et al. Epidemiology of an outbreak of head lice in Georgia. *Am J Trop Med Hyg.* 1976;25(5):739–743.
2. Juranek DD. Epidemiologic investigations of pediculosis capitis in school children. In: Orkin M, ed. *Scabies and Pediculosis.* Philadelphia, Pa, and Toronto, Canada: JR Lippincott Co; 1977:168–173.
3. Slonka GF, Fleissner ML, Berlin J, Puleo J, Harrod EK, Schultz MG. An epidemic of pediculosis capitis. *J Parasitol.* 1977;63(2):377–383.
4. Taplin D, Meinking TL, Castillero PM, Sanchez R. Permethrin 1% creme rinse for the treatment of *Pediculus humanus* var. *capitis* infestation. *Pediatr Dermatol.* 1986;3(4):344–348.
5. Brandenburg K, Deinard AS, DiNapoli J, Englender SJ, Orthoefer J, Wagner D. One

percent permethrin creme rinse vs 1% lindane shampoo in treating pediculosis capitis. *Am J Dis Child.* 1986;140:894–896.
6. Bowerman JG, Gomez MP, Austin RD, Wold DE. Comparative study of permethrin 1% creme rinse and lindane shampoo for the treatment of head lice. *Pediatr Infect Dis J.* 1987;6:252–255.
7. Strom BL. Position paper on the use of purported post-marketing drug surveillance studies for promotional purposes. *Pharmacol Ther.* 1990;48(5):598.

Call for Papers for Symposium on Analysis of Soils Contaminated with Petroleum Constituents

Original papers are needed for the Symposium on Analysis of Soils Contaminated with Petroleum Constituents, sponsored by the American Society for Testing and Materials (ASTM) Committee D-18 on Soil and Rock, in cooperation with the American Petroleum Institute. The symposium will be held June 24 to 25, 1993, in Atlanta, Ga, and will examine the existing technology used by scientists and engineers to measure and evaluate soil material that has become contaminated with organic derivatives of petroleum hydrocarbons or unrefined petroleum.

Papers that address these topics are solicited:

- Contaminated soil sample collection
- Volatile organic compound analysis of soil contaminated with petroleum constituents
- Semi-volatile organic compound analysis of soil contaminated with petroleum constituents
- Pesticide and PCB analysis of soil contaminated with petroleum constituents
- "Total" preparation/analyses of soil contaminated with petroleum constituents: total petroleum hydrocarbons, total organic carbon, total halogenated carbon, oil, and grease
- Contaminated soil data interpretation and statistical analysis

Prospective authors are asked to submit a title, a 250- to 300-word abstract, and an ASTM paper submittal form by *June 23, 1992,* to Dorothy Savini, Symposia Operations, ASTM, 1916 Race Street, Philadelphia, PA 19103-1187; tel. (215) 299-5413. Sending your abstract by fax is not encouraged because copies do not reproduce well. Paper submittal forms are available from Ms. Savini or from the symposium chairman.

A special technical publication based on the symposium proceedings is anticipated by ASTM. For more information contact Symposium Chairman Tracey O'Shay, Gordon and Lawton, P.O. Box 80072, Austin, TX 78727-0072; tel. (512) 475-2302; or Co-Chairman Keith Hoddinott, US Army Environmental Hygiene Agency, attn: HSHB-ME-SR, Aberdeen Proving Grounds, MD 21010-5422; tel. (301) 671-2953.

Committee D-18 is one of 132 ASTM technical standards-writing committees. Organized in 1898, ASTM is one of the largest voluntary standards development systems in the world.

Public Health Briefs

First-Year Impact of the 1989 California Cigarette Tax Increase on Cigarette Consumption

Robert L. Flewelling, PhD, Erin Kenney, PhD, MPH, John P. Elder, PhD, MPH, John Pierce, PhD, Michael Johnson, PhD, and Dileep G. Bal, MD

ABSTRACT

We employed a time series design to evaluate the impact of the 1989 California cigarette tax increase on cigarette consumption in California. Adult per capita consumption data from 1980 to 1990 were analyzed for California and the United States. Trend data indicated a sharp drop in California cigarette consumption coincident with the tax increase. Time-series regression analyses support this observation, and suggest that a 5% to 7% decline in consumption is attributable to the tax increase. (*Am J Public Health.* 1992;82:867–869)

Introduction

With the passage of the voter-initiated California Proposition 99, the state excise tax on cigarettes was increased by 25 cents per package, effective January 1, 1989.[1] Although the initiative earmarked a percentage of the new revenues for tobacco use control, medical care, and tobacco-related research,[2] excise taxes have also been recognized as having potential public health benefits by reducing cigarette consumption.[3–7] The purpose of this study was to estimate the impact of the 1989 tax increase on cigarette consumption in California.

Methods

Cigarette consumption data were derived from state government records of cigarette excise tax revenues.[8] Quarterly adult per capita consumption was computed by dividing the total number of packs of cigarettes taxed by the estimated population over age 17. Quarterly estimates of the adult population were interpolated from yearly estimates published by the US Census Bureau.[9,10] Quarterly prices for cigarettes were interpolated from annual (November) estimates[11] and adjusted for inflation.

A quasi-experimental multiple time series design was used to analyze consumption data from 1980 through the end of 1990. This period offers a relatively stable baseline against which extraneous effects can be quantified. Consumption data from California and the United States were first plotted to visually assess any change in the consumption trends. Time-series regression analyses were conducted to assess the size and significance of the impact of the tax increase. The regression models controlled for the influence of changes in the real price of cigarettes, and for other factors known to influence cigarette distribution and sales. Because of multicollinearity among the predictor variables, three alternative regression models were used: (1) an ordinary least squares (OLS) model with all relevant variables included; (2) an OLS model with an extraneous, prespecified coefficient for the real price of cigarettes; and (3) a ridge regression model with all relevant variables included. The extraneous price coefficient in the second model was based on a −0.4 price elasticity of demand for cigarettes and for consumption and price levels just prior to the tax increase.

Results

Quarterly per capita consumption trends from 1980 to 1990 for California and for all other states combined are plotted in Figure 1. Immediately apparent are the overall lower consumption rate in Califor-

At the time of the study, Robert L. Flewelling was with the Department of Health Behavior and Health Education, School of Public Health, University of North Carolina at Chapel Hill. Erin Kenney and John P. Elder are with the Graduate School of Public Health, San Diego State University. John Pierce is with Population Studies in Cancer Prevention, University of California, San Diego. Michael Johnson and Dileep G. Bal are with the Tobacco Control Section, California Department of Health Services, Sacramento.

Requests for reprints should be sent to Erin Kenney, PhD, MPH, San Diego State University, 6363 Alvarado Court #100, San Diego, CA 92120.

This paper was submitted to the Journal April 23, 1991, and accepted with revisions November 12, 1991.

Editor's Note. See related editorial on p 785 of this issue.

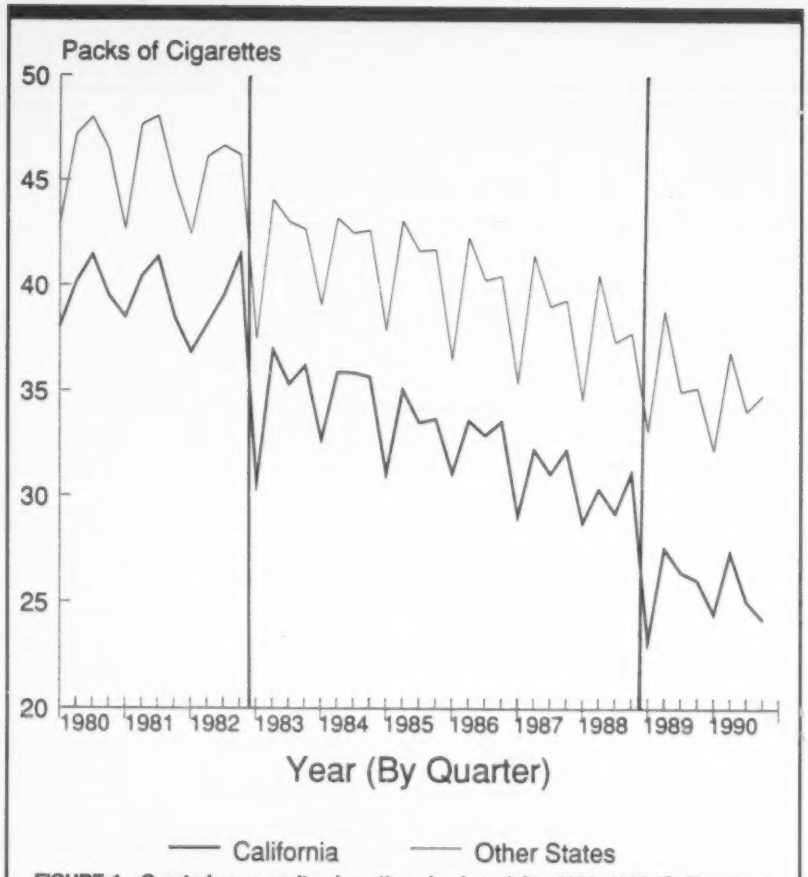


FIGURE 1—Quarterly per capita cigarette sales by adults, 1980–1990 (California vs United States)

TABLE 1—Annual Percentage Decline in Adult per Capita Cigarette Consumption for California and Other States, 1983–1990

	California	Other States
Years before 1989		
1983 to 1984	0.93	1.11
1984 to 1985	4.99	1.86
1985 to 1986	1.56	2.91
1986 to 1987	5.10	2.77
1987 to 1988	4.27	3.16
Mean	3.40	2.44
Year of tax increase		
1988 to 1989	9.30	5.38
Years after 1989		
1989 to 1990	4.30	2.88

Note. Annual percentage declines were computed by averaging the declines for each quarter over the entire year. Quarters immediately before or after the 1983 federal tax increase and (for California) the 1989 tax increase were excluded from calculations.

nia than in all other states, the seasonal patterns by quarter, and a general downward trend in consumption. The 1983 federal tax increase and the 1989 California tax increase are indicated by vertical reference lines. There was a noticeable decline in consumption, which appears to be permanent, at the time of the federal tax increase. As expected, the California tax increase apparently affected consumption at a time exactly coincident with the implementation of the increased tax. Positive pretax and negative posttax consumption spikes, presumably due to stockpiling, are apparent for California and the United States in 1983, and for California in 1989.

The average decline in consumption from 1988 to 1989 was 9.3%, whereas the average yearly decline in the 5 years prior to the tax increase was 3.4% (Table 1). Assuming that the pretax rate of decline would have continued in the absence of the tax increase, the excess drop in consumption following the tax increase may be estimated as the difference between the two rates, or 5.9%. The decline in other states was also somewhat larger than observed in previous years and may be due to sizable tax increases in several other large states.

Table 2 presents the results of the three regression models for predicting quarterly adult per capita consumption in California. The main variable of interest is CALTAX, the dummy variable representing all quarters subsequent to the California tax increase. The coefficients for this variable estimate the overall drop in quarterly per capita consumption coincident with the 1989 California tax increase, when other known effects on cigarette consumption are controlled. The first two models produce very consistent parameter estimates; the third model produces a somewhat larger estimate. Taken together, the results suggest that a drop in per capita consumption of between about 1.5 and 2.0 packs per capita (quarterly) may be attributed to, or is at least coincident with, the 1989 tax increase. This translates to a 5% to 7% drop in per capita consumption from the 1988 level. When similar models were run for all other states combined, no statistically significant effect for CALTAX was found.

Discussion

The present analysis supports the conclusion that the implementation of the 1989 cigarette excise tax increase in California did have a significant impact on adult per capita cigarette consumption as approximated by tax-paid sales. The 5% to 7% decline in consumption derived from the regression models was consistent with the unadjusted trend data. Because the regression models control for the influence of the real price of cigarettes and base the estimated effect on all quarters subsequent to the tax increase, they appear to provide a more valid method for estimating the impact of the tax increase.

An issue that cannot be entirely resolved with the available data concerns the long-term effects of the tax increase. Although the regression models assume a one-step, permanent drop in consump-

tion, the long-term impact of the tax increase is uncertain. Even with additional data, estimating long-term effects is tenuous because of potential confounding forces, such as price, public education, and societal trends, that also influence changes in smoking behavior.

It is possible that the true impact of the California tax increase is slightly lower than the rate of 5% to 7% suggested by this study. Consumption trends in neighboring states indicate that a small percentage of the decline in California may be due to increased interstate bootlegging. Also, although not statistically significant, the unusually large decline in consumption between 1988 and 1989 in other states suggests that other influences besides the tax increase may also have been operative in California.

The estimated impact of the 1989 tax increase represents a price elasticity of between −0.25 and −0.35 and is somewhat less than previous studies on the price elasticity for cigarettes would have predicted. As the social acceptability of smoking continues to wane, smokers may be increasingly resistant to efforts that encourage cessation, including price increases. Thus, a broad array of educational strategies and policy mechanisms will be required to ensure that progress toward a smoke-free society is maintained. (The full report, containing a detailed description of the analytic method and results, is available from the authors upon request.) □

TABLE 2—Regression Coefficients for Predicting Quarterly Adult per Capita Cigarette Consumption in California and Other States, 1980–1990

Variable[a]	Model: Full	Model: Extraneous REALPC[b]	Model: Ridge[c]
CALTAX[d]	−1.46**	−1.50**	2.08
REALPC	−0.11*	(−0.10)	−0.15
UAR	−0.24***	−0.25***	−0.15
Q2DUM	2.86***	2.86***	2.21
Q3DUM	2.62***	2.62***	1.53
Q4DUM	2.53***	2.53***	1.27
PREFED	3.86***	3.85***	3.64
POSTFED	−3.28**	−3.36***	−2.70
PRETAX	1.40	1.39	0.49
POSTTAX	−2.22*	−2.18*	−2.86
R^2	.96	.97	...
Durbin-Watson	2.53	2.52	...

[a]CALTAX is the dummy variable indicator for all quarters since January 1, 1988; REALPC is the estimated real price of cigarettes (excluding the 25-cent increase due to the 1989 tax increase); QUARTER is the quarter number (starting with the first quarter of 1980), used to capture the overall downward linear trend in consumption; Q2DUM, Q3DUM, and Q4DUM are dummy indicators for second, third, and fourth quarter of each year to control for seasonal variation in wholesale distributions; and PREFED, POSTFED, PRETAX, and POSTTAX are dummy variables to capture the pre- and posttax increase sales spikes associated with the 1983 federal and 1989 state excise tax increases.
[b]Extraneous REALPC model is an OLS model with an extraneous, prespecified coefficient for the real price of cigarettes (based on −0.4 price elasticity and for consumption and price levels just prior to the tax increase).
[c]Parameter estimates for $k = .30$.
[d]For each model, the coefficient for the variable CALTAX reflects the average sustained change in quarterly adult per capita consumption associated with the 1989 tax increase.
*$P < .05$; **$P < .01$; ***$P < .001$. (Significance levels not available for ridge regression models.)

References

1. California Revenue and Taxation Code, division 2, part 3, chapter 2, article 2, 30121–30130.
2. Bal DG, Kizer KW, Felton PG, Mozar HN, Niemeyer, D. Reducing tobacco consumption in California: development of a statewide anti-tobacco use campaign. *JAMA.* 1990;264:1570–1574.
3. Warner KE. Cigarette taxation: Doing good by doing well. *J Public Health Policy.* September 1984:312–319.
4. Warner KE, Murt HA. Economic incentives for health. *Ann Rev Public Health.* 1984;5:107–133.
5. Cummings KM, Sciandra R. The public health benefit of increasing tobacco taxes in New York State. *NY State J Med.* 1990;90:174–175.
6. Canadian Coalition on Smoking and Health. *Tobacco Tax Policy in Canada: A Health Perspective.* Ottawa, Ontario, Canada: National Campaign for Action on Tobacco; January 1990.
7. Lewit E, Coate D. The potential for using excise taxes to reduce smoking. *J Health Econ.* 1982;1:121–145.
8. Tobacco Institute. *Cigarette Tax Data—Monthly State Cigarette Tax Report.* Washington, DC: Tobacco Institute; January 1980 through September 1990.
9. IMSL Inc. *Math Library User's Manual.* Vol. 2. Houston, Tex: IMSL; April 1987.
10. US Department of Commerce. *Statistical Abstract of the United States.* Washington, DC: US Bureau of the Census; 1981 through 1990.
11. Tobacco Institute. *The Tax Burden on Tobacco, Historical Compilation.* Vol. 24 (1989). Washington, DC: Tobacco Institute; 1990.

Drug Use and Suicide Ideation and Behavior among North Carolina Public School Students

W. Michael Felts, PhD, Thomas Chenier, PhD, and Robert Barnes, EdD

ABSTRACT

We used data from 3064 respondents to the 1990 North Carolina Youth Risk Behavior Survey to examine the relationship between adolescent drug use and suicide ideation and behavior. Principal components analysis followed by varimax rotation was performed separately on the drug use and suicide items. Correlation coefficients computed between the two sets of factor scores and comparison of mean drug use factor scores revealed that drug use, particularly of crack/cocaine, was related to increased suicide ideation and behavior. (*Am J Public Health.* 1992; 82:870–872)

Introduction

Suicide, now the second leading cause of death among US youth aged 15 to 19,[1] is a preeminent concern among health professionals seeking to improve the health status of adolescents in the United States.[2] The design of effective intervention strategies to address this problem depends on the identification of risk factors related to suicide ideation and behavior. Among the risk factors currently acknowledged to be in some way linked to adolescent suicide are previous suicide attempts, illness, family violence, precipitous life events, and substance abuse.[3]

The purpose of this study was to determine if a relationship existed between suicide ideation and behavior and substance abuse among adolescents attending North Carolina public secondary schools and, if so, to delineate the nature of that relationship.

Methods

These data were collected as part of a larger study of health risk behaviors of North Carolina public school students.[4] The Youth Risk Behavior Survey (YRBS)[5] was administered during May 1990 to students in 9th- and 12th-grade classes that had been randomly selected using a self-weighting two-stage cluster sampling procedure. The survey instrument consisted of 61 forced-choice items, 4 of which addressed demographic questions; the remaining 57 items dealt with student behavior related to (1) violence; (2) tobacco, alcohol, and drug use; (3) exercise; (4) nutrition; and (5) intentional and unintentional injury.

The responses were randomly divided into two subsamples. For each subsample, responses to 13 questions dealing with drug use and 4 questions about suicide ideation and behavior were analyzed. Descriptive statistics were generated for the 17 questions of interest, and principal component factor analysis, with varimax rotation, was applied separately to the drug use and suicide questions.[6] Relationships between the extracted factors were then determined through correlational analysis and comparisons of mean factor scores. Because the results of the cross-validation analyses were identical, all data were pooled to generate the reported results.

Results

The validity of the self-report data was substantiated by comparison with results of recent surveys of comparable populations. Drug use data were compared with those collected approximately 1 year earlier in the statewide Alcohol and Drug Defense (ADD) survey.[7] Because no previous statewide survey of suicide ideation and behavior had been conducted, these results were compared with those of the National Adolescent Student Health Survey.[8] Although these comparisons were imprecise because of differences in item wording and grade range groupings from which data were collected, the North Carolina YRBS (NCYRBS) data closely paralleled the results of the two comparison surveys. For example, 2.3% of the students in grades 7 through 12 reported having used cocaine within the previous 30 days in the ADD survey, compared with 2.4% of 9th and 12th graders in the NCYRBS. In the ADD survey, 13.6% of the students in grades 7 through 12 reported having used marijuana within the previous 30 days; this response was given by 11.0% and 14.0% of 9th and 12th graders, respectively, in the NCYRBS. Thirty-three percent of the NCYRBS respondents reported having thought about suicide and 4.5% reported that they had actually made a suicide attempt in the 12

W. Michael Felts and Robert Barnes are with the Department of Health Education and Thomas Chenier is with the Biostatistics Research Program, East Carolina University, Greenville, NC.

Requests for reprints should be sent to W. Michael Felts, PhD, Department of Health Education, East Carolina University, Greenville, NC 27858.

This paper was submitted to the Journal May 30, 1991, and accepted with revisions November 12, 1991.

TABLE 1—Results of Principal Components Analysis with Varimax Rotation of Response to Thirteen Drug Use Items

Question	Factor 1 Cocaine/Crack	Factor 2 Alcohol	Factor 3 Marijuana	Factor 4 Needle Drugs
During your life, on how many occasions have you used cocaine in any form (including powder, crack, or freebase)?	.91	...	...	...
During the past 30 days, on how many occasions have you used cocaine in any form (including powder, crack, or freebase)?	.91	...	...	...
How old were you when you used cocaine in any form (including powder, crack, or freebase) for the first time?	−.74	...	...	...
During your life, on how many occasions have you used the *crack* or *freebase* form of cocaine?	.88	...	...	...
During the past 30 days, on how many occasions did you drink alcohol (including wine, wine coolers, beer, liquor, or mixed drinks)?	...	.88	...	...
During the past 30 days, how many times did you have five or more drinks on one occasion? (A "drink" is a glass of wine, a wine cooler, a bottle or can of beer, a shot glass of liquor, or one mixed drink.)	...	.81	...	...
During your life, on how many occasions did you drink alcohol (including wine, wine coolers, beer, liquor, or mixed drinks)?	...	.85	...	
How old were you when you had your first drink of alcohol (other than a sip)?	...	−.77	...	...
During your life, on how many occasions have you used marijuana?	...	...	.86	...
During the past 30 days, on how many occasions have you used marijuana?	...	...	.65	...
How old were you when you used marijuana for the first time?	...	...	−.81	...
During your life, on how many occasions have you used any other type of illegal drugs such as LSD, PCP, MDMA, "Ecstasy," mushrooms, speed, or heroin?	...	...	.58	...
During your life, have you ever injected (shot up) any drug not prescribed by a doctor, such as steroids, cocaine, amphetamines, or heroin?	...	...	...	.92

Note. Ellipses indicate that factor loading is less than 0.5 in magnitude.

months prior to survey administration. Approximately 30% of respondents (aged 12 to 19) to the NASHS reported to have seriously thought about committing suicide, with 18% of girls and 11% of boys reporting to have actually attempted suicide.[8] Detailed tables of the North Carolina YRBS data are available from the authors upon request.

Table 1 contains results of the exploratory factor analysis of the 13 drug use items. On the basis of observed eigenvalues, four factors were extracted by means of a principal components algorithm and then subjected to varimax rotation. Factor 1 (cocaine/crack) is made up of four questions dealing with cocaine and crack use. Factor 2 (alcohol) comprises four questions dealing with patterns of alcohol use. Factor 3 (marijuana) is composed of three questions dealing with marijuana use and a single item dealing with the use of a variety of illicit substances including LSD, heroin, and mushrooms. Factor 4 (needle drugs) consisted of a single question dealing with use of injectable drugs. Factor analysis of the four suicide questions yielded a two-factor solution (Table 2). Factor 1 is associated with suicide ideation and intention and factor 2 with suicide behavior and the severity of its outcome.

TABLE 2—Results of Principal Components Analysis with Varimax Rotation of Responses to Suicide Ideation/Behavior Items

Question	Factor 1 Suicide Ideation	Factor 2 Suicide Behavior
During the past 12 months, have you ever *seriously* thought about attempting suicide?	.89	...
During the past 12 months, did you make a specific plan about how you would attempt suicide?	.84	...
During the past 12 months, how many times did you actually make a suicide attempt?	...	.76
If you attempted suicide during the past 12 months, did that attempt result in an injury or poisoning that had to be treated by a doctor or nurse?	...	.93

Note. Ellipses indicate that factor loading is less than 0.5 in magnitude.

Correlation coefficients computed between the drug factor and suicide factors were as follows: suicide ideation/intention with cocaine/crack, $r = -.08$ ($P = .0001$); with alcohol, $r = -.16$ ($P = .0001$); with marijuana, $r = -.10$ ($P = .0001$); with needle drugs, $r = .08$ ($P = .0001$); and suicide behavior/severity with cocaine/crack, $r = .32$ ($P = .0001$); with alcohol, $r = .08$ ($P = .0001$); with marijuana, $r = .11$ ($P = .0001$); with needle drugs, $r = -.07$ ($P = .0003$). The valence of these coefficients is a function of the directionality of the questions loading on each factor. In each case the valence is consistent with the hypothesis that increased drug use,

TABLE 3—Mean Drug Use Factor Scores for Three Suicide Questions

Question[a]		n	Cocaine/Crack[b]	Alcohol[b]	Marijuana[b]	Needle Drugs[c]
During the past 12 months, have you ever *seriously* thought about attempting suicide?	Yes	712	0.17[d]	0.28[d]	0.17[d]	−0.11[d]
	No	2148	−0.05	−0.09	−0.06	0.04
During the past 12 months, did you make a specific plan about how you would attempt suicide?	Yes	417	0.40[d]	0.35[d]	0.25[d]	−0.23[d]
	No	2473	−0.07	−0.06	−0.04	0.04
If you attempted suicide during the past 12 months, did that attempt result in an injury or poisoning that had to be treated by a doctor or nurse?	Yes	57	1.93[d]	0.66[d]	0.82[d]	−0.41
	No	2675	−0.04	−0.03	−0.01	0.02

[a]These questions were prefaced by the following statement: Some people sometimes feel so depressed and hopeless that they may even consider attempting suicide, that is taking some action to end their own life.
[b]Higher scores for the cocaine, alcohol, and marijuana factors represent greater frequency of use.
[c]Lower scores for the needle drugs factor represent greater frequency of use.
[d]Differences in mean scores are significantly different ($P < .01$) in these cells.

early onset of drug use, or both is associated with a greater tendency to think about or actually attempt suicide. This pattern of relationships was supported by comparison of mean drug factor scores as a function of responses to the individual suicide questions. Significant differences existed between mean factor scores for all but one case (needle drug factor and severity of suicide attempt) (Table 3). Correlation coefficients calculated between drug factor scores and the number of reported suicide attempts were as follows: cocaine/crack, $r = .29$ ($P = .0001$); alcohol, $r = .13$ ($P = .0001$); marijuana, $r = .11$ ($P = .0001$); and needle drugs, $r = -.09$ ($P = .0001$).

Discussion

The relationship between substance abuse and suicide ideation and behavior has been explored by a number of researchers. Substance abuse appears to be associated with more frequent suicide ideation and suicide attempts, particularly in retrospective studies among high-risk youth.[9] Our study generalizes these relationships to noninstitutionalized public school youth and further specifies the nature of the substance abuse–suicide relationship.

In this study the use of cocaine/crack was more closely associated with a self-reported incidence of attempted suicide than was use of alcohol, marijuana, or needle drugs. This effect may have been due to a differential influence produced by the substance of choice or to a tendency of individuals with a heightened susceptibility to the emotional conditions leading to attempted suicide to use cocaine/crack. Use of cocaine/crack, alcohol, or marijuana was significantly related to a student's report of seriously thinking about attempting suicide or making specific suicide plans.

A precise cause-and-effect relationship between drug use and suicide behavior cannot be established from these data. Information concerning previous life events, family violence, illness, and other potential confounding factors was not available in the NCYRBS data base. The results of this study do, however, agree with the findings of previous researchers linking substance abuse to adolescent suicide. Recognition of this relationship is important for health educators involved in designing intervention programs for high-risk youth as well as for the general school population. □

Acknowledgments

This research was completed with support provided by a cooperative agreement (U63/CCU403072) with the Division of Adolescent and School Health, Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control, Atlanta, Ga.

This paper was presented at the annual conference of the Southern District of the American Alliance of Health, Physical Education, Recreation, and Dance, February 16, 1991.

References

1. US Department of Health and Human Services. *Healthy Prevention Objectives.* DHHS Pub. No. (PHS)91-505212. Washington, DC: Govt Printing Office, 1991.
2. National Commission on the Role of the School and the Community in Improving Adolescent Health. *Code Blue: Uniting for Healthier Youth.* Alexandria, Va: National Association of State Boards of Education; 1990.
3. Neiger BL, Hopkins RW. Adolescent suicide characteristic traits of high-risk teenagers. *Adolescence.* 1988;90:468–475.
4. HIV Prevention Education Research Team. *The 1990 North Carolina Youth Risk Behavior Survey.* An unpublished report available from the North Carolina Department of Public Instruction, Healthful Living Section, 116 W Edenton Street, Raleigh, NC 27603.
5. Kolbe L. AAHE Scholar Address: an epidemiological surveillance system to monitor the prevalence of youth behaviors that most effect health. *Health Educ.* 1990;21:44–49.
6. SAS Institute, Inc. *SAS User's Guide: Statistics, Version 5.* Cary, NC: SAS Institute, Inc; 1985.
7. Alcohol and Drug Defense Section. *Alcohol and Other Drug Use Patterns Among Students in North Carolina Public School Grades 7–12.* An unpublished report available from the North Carolina Department of Public Instruction, Division of Student Services, 116 W Edenton Street, Raleigh, NC 27603.
8. American School Health Association, Association for the Advancement of Health Education, and Society for Public Health Education Inc. *The National Adolescent Student Health Survey: A Report on the Health of America's Youth.* Oakland, Calif: Third Party Publishing; 1989.
9. Crumley FE. Substance abuse and adolescent suicide behavior. *JAMA.* 1990;263:3051–3056.

Delay in Ambulance Dispatch to Road Accidents

ABSTRACT

When a road accident occurs, the police communications officer, or 911 operator, generally receives the first call. If the caller reports injuries, the emergency medical services dispatcher is notified immediately; but if the caller is uncertain of injuries, the operator may wait. Most often an ambulance is not needed. However, in nearly 20% of fatal road accidents in Missouri, waiting for confirmation of need resulted in a delay of 5 minutes or more in the dispatch of an ambulance. (*Am J Public Health.* 1992;82:873–875)

Harold Brodsky, PhD

Introduction

In life-threatening road accidents, a call for emergency medical services (EMS) needs to go out quickly and unambiguously. Trauma is a "time-dependent disease," and basic life support may be needed soon after a crash.[1] Any factor that might delay an EMS arrival at a road accident where an injury has occurred should be a matter of concern.

A police communications officer or 911 operator usually receives the earliest notification of an accident. The police depend on the caller for information on possible injuries. If the caller specifies injuries, the EMS dispatcher is notified quickly and an ambulance may be sent out almost immediately. If the caller is vague, however, the police officer may wait before contacting EMS.

How often does a policy of waiting for confirmation of need result in the delayed dispatch of an ambulance to an injury accident? Many states do not require ambulance services to submit data on notification time, so there is no way of telling. Fortunately, Missouri has police and EMS records that can be used for this type of statistical analysis. However, in other respects, such as accident rate and police communications policy, Missouri appears to be typical of the United States and perhaps other Western nations.

Analysis of Dispatch Delays

Two data sources from Missouri are used here: fatal accidents between 1985 and 1988, and injury accidents in 1987. Details on the processing of these data are discussed elsewhere.[2] Estimates used here for delays are believed to be conservative, especially for injury accidents.

A 4-year record of fatal accidents shows that in 18.2% of the cases (nearly one out of five), EMS was notified 5 minutes or more after the police were called (Table 1). A 1-year summary of injury accidents involving ambulance transports shows that the probability of delay varies with the police-reported severity of injury. Police-reported severities of injury, by and large, are not reliable substitutes for clinical measures, but they can be used as an index of outward appearance. Generally, the less obvious the appearance of an injury, the more likely that dispatch of an ambulance will be delayed (Table 2).

The most reasonable explanation for delays of 5 minutes or more is vague calls to the police about injuries. A caller might be uncertain whether an ambulance is needed if injuries are not apparent. Injuries that are less likely to be apparent are among those classified by the police as less severe (Table 2).

Callers might also be vague about injuries if they did not stop to investigate before calling the police. Variables significantly associated with longer delays are land (urban vs rural), time (normal activity

The author is with the Department of Geography, University of Maryland, in College Park.

Requests for reprints should be sent to Harold Brodsky, PhD, Department of Geography, University of Maryland, College Park, MD 20742.

This paper was submitted to the Journal April 2, 1991, and accepted with revisions November 27, 1991.

TABLE 1—EMS Notification Time by Police: Fatal Accidents, Missouri, 1985–1988

	Accidents (n = 3290)		Cumulative Accidents (n = 3290)	
Notification Time[a]	no.	%	no.	%
1 min or less EMS delay	2218	67.4	2218	67.4
2–4 min EMS delay	474	14.4	2692	81.8
5–50 min EMS delay	598	18.2	3290	100.0

Note. Delays of more than 50 minutes were excluded to avoid extreme examples.
[a]EMS notification time from Federal Fatal Accident Reporting System (FARS) data; police notification time from Missouri Highway Accident File.

TABLE 2—EMS Notification Time by Police-Estimated Severity of Injury: Injury Accidents, Missouri, 1987[a]

Notification Time	Police-Estimated Severity of Injury			
	Fatal (n = 527) %	Disabling (n = 5193) %	Evident (n = 6841) %	Probable (n = 3023) %
1 min or less EMS delay	66.6	66.1	60.9	51.1
2–4 min EMS delay	17.6	17.8	17.6	16.4
5–50 min EMS delay	15.8	16.0	21.5	32.5
Total	100.0	100.0	100.0	100.0

Note. Delays of more than 50 minutes were excluded to avoid extreme examples, and some data are missing due to incomplete or incorrect records for matching.
[a]Taken from matched Missouri Police Accident Data and from EMS Ambulance Data 1987.

TABLE 3—Logit Analysis: Fatal Accidents, Missouri, 1985–1988

Independent Variable or Effect	Parameter	Estimate	Standard Error	Chi-Square	P
Intercept	1	1.22589	.0556978	484.43	.0001
Land (urban 1, rural –1)	2	0.164661	.0556804	8.75	.0031
Time (normal 1, late –1)	3	0.148803	.054566	7.44	.0064
Road (interstate 1,0) (state 0,1) (local –1, –1)	4	–.519816	.0844427	37.89	.0001
	5	0.138632	.066491	4.35	.0371
Land × time	6	0.12065	.0541755	4.96	.0259

Note. Dependent variable is ambulance dispatch delays of 1 minute or less, relative to (or divided by) delays of 5 minutes or more.
[a]A copy of the data set and the control statements to run this analysis under SAS can be obtained from the author.

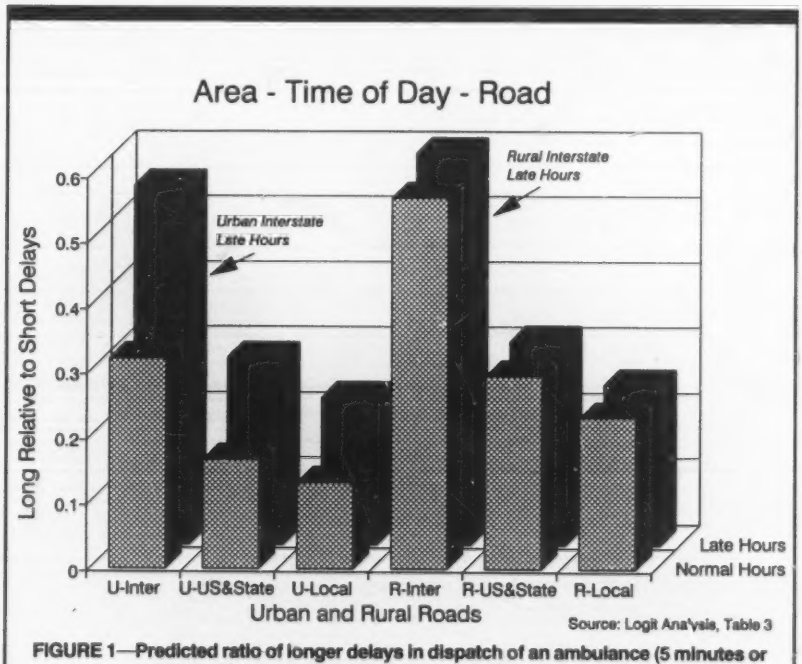


FIGURE 1—Predicted ratio of longer delays in dispatch of an ambulance (5 minutes or more) to shorter delays (1 minute or less) in fatal accidents in Missouri, 1985–1988, by road environment

hours of 7 AM to 11 PM vs late hours), and road (interstates, US and state highways, and local county roads and urban streets). The interaction between land and time indicates that during normal hours of the day, an urban environment has a lower probability of delay than a rural environment, but late at night, urban and rural environments have similar probabilities (Table 3 and Figure 1).

The highest probabilities of delay, regardless of time of day and area, are from the interstate road system. Calls about accidents on interstate highways may come from passersby who see the accidents but do not stop to investigate. A motorist might sense a danger in stopping and backing up on a high-speed interstate highway, especially late at night. The first call to the police, therefore, might not be specific about injuries. In contrast, if accidents occur on roads passing through populated areas, bystanders or motorists will usually check to see if there are injuries before calling. Nevertheless, late at night, even in an urban area (as the interaction variable seems to indicate), the first caller may report only having heard a loud crash.

Discussion

Police communications officers are currently the gatekeepers of emergency services in road accidents.[3] They are required to estimate the likelihood of injuries when a vague call is received. Most often an ambulance is not needed. The chance of any single road accident being life-threatening is small, and so the police may perceive that the risk involved in waiting for confirmation of need is also very small. Yet in almost one out of five fatal accidents in Missouri, the police were initially unaware of the severity of the injury, and a delay of 5 minutes or more occurred before they notified EMS.

Police communication policy on handling vague calls varies with locale. In some smaller cities in Missouri, EMS is apparently notified in all, or almost all, local road accidents—perhaps because it usually takes little time for an ambulance to be sent to the scene. In certain counties, the police officers notify EMS immediately if the crash occurs in a remote rural location, even if the caller is vague about injuries. These officers reason that if they wait for confirmation of need, it might take too long and EMS might arrive too late to do any good.

Some EMS managers in Missouri feel that in an uncertain situation, these basic decisions should be made by the

EMS dispatcher and not the police communications officer. EMS has several options. If an ambulance is available and another is in reserve, EMS may prefer to go out to a road accident even if the probability of need seems small. The ambulance could proceed at normal speed, without the use of emergency sirens, and remain alert to a probable recall.

Some say that ambulances could be dispatched more often when injuries are uncertain. Such a policy would reduce the number of communication delays in life-threatening road accidents. But at what cost? Apparently there are differences in opinion on this matter.

Because conditions vary locally, no blanket policy recommendation may be possible. At present, communications officers have little feedback on the overall consequences of their decisions. However, it is possible to link police communication data with emergency medical data, which makes it possible to monitor the effects of any change in policy on outcome and cost.

No one knows for sure how many lives could be saved or recoveries eased by the faster arrival of an ambulance. But no one will argue against the need for a rapid EMS response when lives may be in danger. □

Acknowledgments

This paper is based on a final report to the AAA Foundation for Traffic Safety in Washington, DC, which supported this research with a grant. Data were obtained with the cooperation of the Missouri State Highway Patrol and the Missouri Department of Health. I would like to thank the Missouri chapter of the Associated Public-Safety Communications Officers, Inc., for an invitation to discuss communication problems at one of their meetings.

Any errors in this paper are my responsibility.

References

1. Stewart RD. Prehospital care of trauma. In: McMurtry RY, McLellan BA, eds. *Management of Blunt Trauma.* Baltimore, Md: Williams & Wilkins; 1990:23–29.
2. Brodsky H. *Geographic Perspectives on Improving Emergency Notification in Road Accidents.* Washington, DC: AAA Foundation for Traffic Safety; 1990.
3. Scott EJ, Percy SL. Gatekeeping police services: police operators and dispatchers. In: Bennett RR, ed. *Police at Work: Policy Issues and Analysis.* Beverly Hills, Calif.: Sage Publications; 1983:127–144.

Race, Belief in Destiny, and Seat Belt Usage: A Pilot Study

Israel Colón, PhD

ABSTRACT

A survey of 1063 individuals found that when belief in destiny was statistically controlled, differences in seat belt use by race disappeared. Thus, racial differences in seat belt use are statistically accounted for and might be explained by belief in destiny. Efforts to increase seat belt use should target minority groups rather than include them in broadbrush programs. Further, these efforts should take into account this important difference in motivation. (*Am J Public Health.* 1992;82:875–877)

Introduction

In the effort to increase seat belt usage, numerous strategies (legislation, educational campaigns, monetary incentives, etc.) have been employed with varying degrees of success. Of particular concern are racial and ethnic minority groups, who have lower rates of seat belt usage than Whites.[1,2] Additionally, Blacks have been found to be less responsive to efforts to increase use.[2]

Race per se (i.e., a physical attribute) certainly cannot explain differences in seat belt usage. One consideration that might be taken into account in targeting minority groups is differences in motivation. One possible difference in motivation is adherence to the concept of destiny. If one is strongly wed to the notion that the end of life is divinely decreed, efforts based on reason or logic will fall on deaf ears. Rotter[3] found that minority groups—Blacks and Puerto Ricans—were more fatalistic than Whites. In a study of seat belt use in families, Williams[4] found correlations between a measure of destiny and seat belt use by mothers and daughters. In another study, seat belt users identified themselves as fatalistic less often than nonusers.[5]

The purpose of this study is to assess the extent to which belief in destiny might explain lower seat belt usage among racial and ethnic minority groups.

Method

The study was carried out during February and March of 1988 in the city of Harrisburg, Pennsylvania, and its suburbs.

In line with the aim of quickly surveying a large number of people who were not being paid to participate in the study, a data entry sheet was developed resembling a data file (i.e., variables were represented in the columns and cases in the rows). Responses were directly entered into these data sheets. Thus, an interview could be completed in a matter of minutes.

A combination availability and purposive sample design was used. The greater Harrisburg area contains only

The author is with the School of Social Administration, Temple University, Philadelphia, PA.

Requests for reprints should be sent to Israel Colón, PhD, 4439 Sansom Street, Philadelphia, PA 19104.

This paper was submitted to the Journal July 22, 1991, and accepted with revisions November 12, 1991.

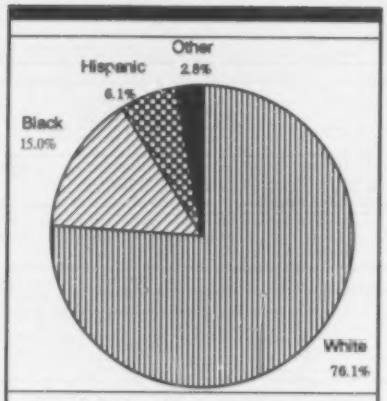


FIGURE 1—Distribution of race (n = 1063)

7.3% Blacks and Hispanics, and the purpose of this study dictated that these people be overrepresented in relation to their relative size in the population.[6] A probability sampling technique, short of using an extremely large sample, would not produce a sufficient number of minority group members to facilitate analyses. Consequently, two criteria were used in selecting participants: (1) maximization of variation—i.e., subjects were sought in a wide array of settings (shopping malls, a motor vehicle bureau, public and private offices, etc.); and (2) overrepresentation of minority group members—i.e., minorities were visually sought out for inclusion and their communities were canvassed.

In addition to demographic information, participants were asked whether they owned an automobile with seat belts that automatically fastened. Anyone who answered yes to this question was precluded from the study. Seat belt usage was operationalized by asking people to approximate the percentage of the time they used seat belts. To measure destiny, the following statement was read to participants: "There is no point in using seat belts since you can't change your destiny." Participants were given a 5 × 8-in card with a seven-point Likert scale, which ranged from *strongly disagree* to *strongly agree,* and were then asked to report their level of agreement with the statement.

Results

The sample was made up of 1063 participants: 690 females (64.9%) and 373 males (35.1%). The mean age of the sample was 38.3 years (SD = 17.1 years). The modal category for race was *White* (76.1%), with minority group members making up 23.9% of the sample (see Figure 1). The mean for seat belt usage ("what percent of the time do you use seat belts?") was 73.6% (SD = 34.2%), with a range running the full gamut (0% to 100%). In response to the statement, "There is no point in using seat belts since you can't change your destiny," the mean level of agreement was 2.2 (SD = 1.7), which corresponds to *disagree* on the seven-point Likert scale.

For purposes of data analysis, the variable for race was dichotomized into *White* and *non-White* categories. Non-Whites' level of agreement (mean = 2.46, SD = 1.9) was significantly greater than that of Whites (mean = 2.08, SD = 1.6) regarding the seat belt–destiny statement (t = 3.14, df = 1061, P = .0017). Non-White seat belt usage (mean = 66.8%, SD = 35.2%) was significantly and substantially lower than that of Whites (mean = 75.4%, SD = 34.1%).

To examine the relationship between race and seat belt usage while statistically controlling for belief in destiny, a two-factor analysis of variance model was used. In this analysis, the seat belt–destiny variable was collapsed into three categories (disagree, neutral, and agree). A summary of this analysis reveals that there is no significant interactional effect and that the only significant predictor of seat belt use is the destiny variable (see Table 1). It is thus apparent that the original difference by race in seat belt usage could be an artifact that disappears once belief in destiny is statistically taken into account.

TABLE 1—Summary of Two-Factor Analysis of Variance Predicting Seat Belt Use

Variable	df	Sum of Squares	Mean Square	F Ratio	P
Destiny	2	110133.29	55066.65	52.60	.0001
Race	1	1136.62	1136.62	1.09	.296
Destiny × race	2	3458.47	1729.24	1.65	.193
Error	1057	1107483.34	1047.76		

Discussion

The purpose of this study was to examine the relationship between race and seat belt usage vis-à-vis the belief in destiny. A multivariate model was used to examine simultaneously the contributions of race and belief in destiny in explaining reported seat belt usage. Differences in usage by race were statistically accounted for by differences in the belief in destiny.

Two limitations of this study are rather apparent: a nonprobability sample was used and seat belt usage was based on self-report. It has been demonstrated that reported use of seat belts is greatly inflated in comparison with observed use.[7,8] Given the exploratory nature of this study, reported use should be taken as an index rather than an estimate of actual use. Because this study was not intended to get actual use estimates but to explore the hypothesis relating race, belief in destiny, and usage, the present findings are certainly provocative and warrant further and more careful study.

Although this study is only exploratory, its findings suggest two points with implications for social policy. First, broadbrush strategies that do not take into account racial differences in motivation to use seat belts can have only limited success. Second, interventions based on logic or rationality appear to be inappropriate strategies with Hispanics and Blacks. Instead, programs might be developed in conjunction with churches in Hispanic and Black communities, where the church has an important voice. Although some may view such an approach as alien to American tradition and values, public safety and welfare suggest pragmatism over idealism. □

References

1. Council FM. *Seat Belts: A Follow-up Study of Their Use under Normal Driving Conditions.* Chapel Hill, NC: University of North Carolina Highway Safety Research Center; 1969.
2. Hunter WW, Campbell BJ, Stewart JR. Seat belts pay off: the evaluation of a community-wide incentive program. *J Safety Res.* 1986;17:23–31.
3. Rotter JB. Generalized expectations for internal versus external control of reinforcement. *Psychol Monogr.* 1966;80:entire issue.
4. Williams AF. Factors associated with seat belt use in families. *J Safety Res.* 1972;4(3): 133–139.
5. US Department of Transportation, National Highway Traffic Safety Administration. *Motivating Factors in the Use of Restraint Systems.* Final report. Philadelphia, Pa: National Analysts, Inc; September 1971.

6. US Bureau of the Census. *State and Metropolitan Area Data Book, 1986.* Washington, DC: US Govt Printing Office; 1986.
7. Waller PF, Barry PZ. *Seat belts: a comparison of observed and reported use.* Chapel Hill, NC: University of North Carolina Highway Safety Research Center; 1969.
8. Market Opinion Research, General Motors Corporation. *An Analysis of the Factors Affecting Seat Belt Use.* Detroit, Mich: Market Opinion Research, General Motors Corporation; 1977.

Work-Site Nutrition Intervention and Employees' Dietary Habits: The Treatwell Program

Glorian Sorensen, PhD, MPH, Diane M. Morris, PhD, RD, Mary K. Hunt, MPH, RD, James R. Hebert, ScD, Donald R. Harris, PhD, Anne Stoddard, ScD, and Judith K. Ockene, PhD

ABSTRACT

In a randomized, controlled study of the Treatwell work-site nutrition intervention program, which focused on promoting eating patterns low in fat and high in fiber, 16 work sites from Massachusetts and Rhode Island were recruited to participate and randomly assigned to either an intervention or a control condition. The intervention included direct education and environmental programming tailored to each work site; control work sites received no intervention. A cohort of workers randomly sampled from each site was surveyed both prior to and following the intervention. Dietary patterns were assessed using a semiquantitative food frequency questionnaire. Adjusting for work site, the decrease in mean dietary fat intake was 1.1% of total calories more in intervention sites than in control sites ($P < .005$). Mean changes in dietary fiber intake between intervention and control sites did not differ. This study provides evidence that a work-site nutrition intervention program can effectively influence the dietary habits of workers. (*Am J Public Health.* 1992;82:877–880)

Introduction

This paper presents the results of a randomized study testing the efficacy of Treatwell, a work-site nutrition intervention program designed to promote dietary changes associated with the reduction of cancer risk. This program focuses on two dietary goals of the National Cancer Institute: reducing consumption of fat to 30% of total calories and increasing consumption of dietary fiber to 20 to 30 g per day.[1] The study tested the hypotheses that there would be a significant decrease in calories from fat and increase in grams of fiber consumed by employees of the intervention work sites compared with employees of control work sites.

Methods

Sixteen work sites from central Massachusetts and Rhode Island were recruited to participate in the study, described elsewhere.[2] Work sites ranged in size from 300 to 1400 employees, and each had a cafeteria with a kitchen. The work sites, stratified by size and distribution of employees by gender, were randomly assigned to an intervention or control condition. At intervention work sites, programs such as classes and food demonstrations were targeted to individuals, and cafeteria point-of-choice labeling programs provided environmental support. In addition, programs were tailored to each work site through an employee advisory board, and eating pattern guidelines provided consistent messages across intervention activities (Appendix 1). No intervention was provided at the control work sites.

A questionnaire was distributed prior to the intervention in summer 1988 (time 1) to 275 employees randomly selected from each work site. Respondents were surveyed a second time at the end of the intervention period (time 2). This survey included a 67-item semiquantitative food frequency questionnaire (FFQ) developed and validated by Willett and colleagues.[3] Employee characteristics, including sex, age, education, ethnicity, occupation, and body mass index (BMI), were also assessed.

One work site assigned to the intervention group elected not to participate in the intervention. Nevertheless, dietary changes of employees of all eight companies assigned to the intervention were compared with those of employees at the control sites. This approach permitted estimation of the overall effectiveness of this program in work sites that agreed to participate, regardless of their "compliance" with the program.

Analyses of the major outcomes were based on repeated measures analysis of variance using the SPSS-X software program.[4] In analyses, employee effects were

Glorian Sorensen and Mary K. Hunt are with the Division of Cancer Epidemiology and Control, Dana-Farber Cancer Institute, Boston, Mass. Diane Morris is with Mainstream Nutrition, Winnipeg, Manitoba. James R. Hebert and Judith K. Ockene are with the Division of Preventive and Behavioral Medicine of the University of Massachusetts Medical School in Worcester. Donald R. Harris and Anne Stoddard are with the University of Massachusetts' School of Public Health in Amherst.

Requests for reprints should be sent to Dr. Glorian Sorensen, Division of Cancer Epidemiology and Control, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115.

This paper was submitted to the Journal February 20, 1991, and accepted with revisions August 22, 1991.

TABLE 1—Comparison of Characteristics Measured at Baseline between Time 1 Survey Nonrespondents and Subjects Completing Both Surveys

Characteristic	Time 1 Survey Nonrespondents (n = 840)	Responders Completing Both Surveys (n = 2011)	*P* Value
Age			
17–29	25.4	21.7	
30–54	60.0	62.9	
55–75	14.6	15.4	.11
Education			
High school graduate	41.1	36.2	
Some college	32.7	29.7	
College graduate	26.2	32.8	<.01
Sex (% male)	47.4	47.6	.46
Mean total fat (% cal.)[a]	34.1 (7.0)	34.0 (6.6)	.46
Mean total dietary fiber (g)[a]	14.55 (7.35)	14.43 (7.16)	.12

[a]Mean (standard deviation).

TABLE 2—Percentage of Employees Indicating That They Participated in Study Activities by Intervention Status

Study Participation Question	Intervention Companies (n = 947)	Control Companies (n = 1064)	*P* Value
Are you aware of a program where you work that teaches employees about nutrition?	67.4	10.5	<.01
Have you attended classes on diet, nutrition, or weight management where you work in last 12 months?			
Baseline	6.1	4.5	.16
Follow-up	16.6	4.4	<.01
Have you participated in any of the following types of events where you work in the last 12 months?			
Cholesterol screening	28.4	28.3	.99
Taste tests of healthy foods	28.9	0.6	<.01
Healthy barbecue	3.3	1.0	<.01
Recipe contest	1.7	0.7	.05
Party with healthy snacks	11.0	2.3	<.01

TABLE 3—Employee Characteristics by Treatment Status[a]

Employee Characteristics	Intervention Companies		Control Companies	
	Mean	SD	Mean	SD
Age, y	39.3	3.41	41.8	3.22
Education, y[b]	14.1	0.57	13.7	1.10
Ethnicity (% White)	94.0	6.01	96.4	2.70
Sex (% male)	43.2	14.93	50.8	21.67
Body Mass Index (mean)[c]				
Males	26.2	0.51	26.6	0.41
Females	24.4	0.66	24.3	0.76

Note. SD = standard deviation.
[a]None of the differences between treatment groups was found to be significant using analysis of variance.
[b]Education was measured on a six-category scale ranging from being less than high school graduate to having an advanced degree. The midpoint of each interval (in years) was used to estimate the mean.
[c]Body mass index = (weight in kg ÷ height in m^2).

nested within work-site effects. Mean values presented for each treatment condition represent the mean of the work-site mean values, thereby giving equal weight to each work site, which served as the unit of randomization and intervention.

Fat was expressed as a percentage of total calories because such a measure provides an intrinsic adjustment for caloric intake and directly addresses the public health message related to percent of calories as fat. It is a conservative measure in that the inflation of the numerator (i.e., fat calories) inflates the denominator by a numerically equivalent amount. Fiber values were transformed by the natural logarithm (ln) to comply with the assumptions of normality and homoscedasticity. Where summary mean scores are presented, we present the geometric mean.[5]

Results

A total of 3076 employees (70%, range: 52% to 93%) responded to the survey at time 1. Of those, 2258 (74%, range: 43% to 90%) returned the survey at time 2. Analyses were ultimately conducted on 2011 respondents for whom adequate data were available from both the time 1 and time 2 surveys.* No differences in baseline fat and fiber consumption, age, or sex were found between those who responded to both surveys and those responding only at time 1; however, those with more education were more likely to return the second survey (Table 1).

Significantly more intervention than control site employees reported that they were aware of a nutrition program at work and that they had attended classes at work on nutrition or weight management (see Table 2).

There were no statistically significant differences in baseline employee characteristics between the intervention and control work sites (see Table 3). Table 4 presents the relationship of these employee characteristics to total intake of dietary fat and dietary fiber (geometric mean), measured at baseline. As shown in the table, total dietary fat varied significantly by age, sex, and BMI; total dietary fiber varied significantly by age, sex, and education. The variability by work sites in mean changes in fat and fiber consumption are shown in Table 5.

There was a treatment-by-time effect for total dietary fat, and the decrease in

*At time 1, 211 questionnaires (7%) were eliminated from the analyses because there were more than 10 blank responses on the FFQ or because the reported daily caloric intake was less than 600 kcal or more than 4000 kcal, suggesting possible reporting errors. Similarly, 123 records were not included in the time 2 analyses for these reasons. Further information about the analyses is available upon request.

TABLE 4—Mean Baseline Level of Dietary Fat and Dietary Fiber by Employee Characteristics and Results of ANOVA Testing of Differences in Means

Employee Characteristic	Mean Total Dietary Fat (% cal.)	F Test (*df*)	*P* Value	Geometric Mean Total Dietary Fiber[a]	F Test (*df*)	*P* Value
Age						
17–35	34.1			12.56		
36–50	35.0			12.41		
51–75	33.3	7.86 (2, 1827)	<.01	14.19	9.96 (2, 1827)	<.01
Education						
High school	34.1			12.16		
Some college	33.9			13.32		
College graduate	34.2	0.37 (2, 1887)	.69	13.46	6.99 (2, 1887)	<.01
Ethnicity						
White	34.2			12.89		
Nonwhite	30.4	0.32 (1, 1908)	.57	16.14	0.01 (1, 1908)	.94
Sex						
Male	33.7			13.47		
Female	34.4	4.28 (1, 1901)	.04	12.37	12.15 (1, 1901)	<.01
Body Mass Index						
Underweight	34.2			11.95		
Acceptable weight	33.8			13.07		
Moderately overweight	33.9			13.13		
Obese	35.5			13.35		
Morbidly obese	37.6	2.50 (4, 1829)	.04	12.09	1.32 (4, 1829)	.26

Note. ANOVA = analysis of variance; *df* = degrees of freedom.
[a]Geometric mean for total dietary fiber intake derived by exponentiating the log value.

mean dietary fat was greater in intervention than in control work sites (Table 6). The decrease in total dietary fat was not affected by age, education, sex, BMI, or ethnicity (not shown in table). There was suggestion of a larger increase in mean dietary fiber among intervention than among control work sites. However, differences in total dietary fiber (ln) remained nonsignificant despite the inclusion of age, sex, education, BMI, and ethnicity in the model. Table 6 also presents the mean changes in fat and fiber after controlling simultaneously for age, sex, and education—the variables most likely to influence the effect of the intervention.

Discussion

This study is one of the first randomized controlled studies of a comprehensive work-site intervention promoting dietary change to reduce cancer risk.[6] The educational messages emphasized two changes: eat less fat and more fiber. The results indicate that the program was efficacious in reducing fat intake among employees of the intervention work sites.

There are many challenges inherent in work-site-based intervention research. In this study, three intervention sites experienced labor strikes and one moved to another location; similar changes did not occur in any of the control sites. One work site randomized to the intervention withdrew from the program although it did permit administration of the second survey. Thus, the program's overall effect was diluted by sites not fully implementing the program. Finally, the aim of 30% or less of calories from fat was not achieved. Nonetheless, we did observe an intervention effect for fat intake despite fluctuations in the degree to which the intervention was implemented.

TABLE 5—Mean and Standard Deviation for Changes in Dietary Fat and Dietary Fiber by Work Site

Work Site	n	Mean Change in Dietary Fat (% cal.) Mean	SD	Mean Change in Dietary Fiber (ln) Mean	SD	Geometric Mean Dietary Fiber[a]
Intervention companies						
1	78	–2.3	7.0	0.89	1.15	2.44
2	96	–1.2	6.4	1.04	1.19	2.83
3	118	–0.9	6.5	1.20	1.10	3.32
4	100	–2.5	7.7	1.20	0.95	3.32
5	192	–0.9	7.3	1.16	1.08	3.19
6	214	–2.3	6.6	1.24	1.13	3.46
7	73	–1.8	7.5	0.87	1.44	2.39
8	76	–1.6	7.6	1.28	1.06	3.60
Control companies						
1	151	–0.2	7.0	–0.138	1.20	2.86
2	157	–0.6	6.0	0.408	1.18	3.10
3	148	–1.0	6.5	0.756	1.12	3.10
4	78	–0.4	7.1	–0.133	1.11	2.59
5	152	–1.2	6.8	0.099	1.21	–2.83
6	116	0.0	5.8	0.304	1.14	2.46
7	128	–1.2	6.6	–0.495	1.19	1.86
8	134	–0.3	7.4	0.950	1.17	3.10

Note. SD = standard deviation.
[a]Geometric mean for total dietary fiber intake derived by exponentiating the log value.

TABLE 6—Observed and Adjusted Changes in Dietary Fat and Dietary Fiber: Results of a Repeated Measures ANOVA

Outcome Variable	Intervention Companies			Control Companies			Treatment by Time F Test	*P* value (*df* = [1,14])
	Time 1	Time 2	Change	Time 1	Time 2	Change		
Total fat (% cal.):								
Observed means	34.3	32.6	–1.70	33.8	33.2	–0.60	12.41	<.01
Adjusted changes[a]			–1.20			–0.50	9.33	<.01
Total fiber (ln):								
Observed means	2.55	2.58	0.03	2.58	2.60	0.02	0.21	.66
Adjusted changes[a]			0.03			0.01	0.48	.50
Geometric means[b]	12.84	13.26		13.22	13.49			

Note. ANOVA = analysis of variance; *df* = degrees of freedom.
[a]Changes adjusted for differences between intervention groups in age, sex, and education.
[b]Geometric mean for total dietary fiber intake derived by exponentiating the log value.

The FFQ was used in this study because it can be self-administered, can be processed relatively rapidly and at a low cost, and can provide estimates of specific nutrient levels. However, it is subject to recall bias because respondents must average their intake of specified foods over a 12-month period. Assuming there is no intervention-related measurement bias, the FFQ can provide a reliable estimate of nutrient levels to be compared within this study between baseline and follow-up. Because the FFQ is associated with a bias in nutrient intake estimates, the results cannot be compared reliably with estimates derived from other types of dietary assessments. The lack of a significant fiber effect may also be a function of the measurement limitations, given that the FFQ may not provide a sensitive assessment of dietary fiber intake.[7]

In conclusion, this study provides some of the first evidence from a randomized controlled trial that a nutrition intervention program in work sites can have a significant impact on workers' eating behaviors. This employee-driven intervention program integrated individual behavioral change strategies with social and environmental support to produce dietary changes that are associated with reduced cancer risk. With growing evidence of the health implications of dietary habits, such strategies must be identified to promote healthy eating patterns not only among those at high risk but also among the entire population. □

Acknowledgments

This study was funded by the National Cancer Institute, Grant 5R01 CA46028.

The authors wish to thank the 16 participating work sites that made this study possible. The authors are also grateful to Dr. Karen Glanz and two anonymous reviewers for their helpful comments on earlier versions of this manuscript.

References

1. Butrum RR, Clifford CK, Lanza E. NCI dietary guidelines: rationale. *Am J Clin Nutr.* 1988;48:888–895.
2. Sorensen G, Hunt MK, Morris DH, et al. Promoting healthy eating patterns in the worksite: the Treatwell intervention model. *Health Educ Res.* 1990;5:505–515.
3. Willett WC, Sampson L, Stampfer MJ, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol.* 1985;122:51–65.
4. *SPSS-X Users Guide.* 2nd ed. New York, NY: McGraw Hill; 1986.
5. Snedecor GW, Cochran WG. *Statistical Methods.* Ames, Iowa: Iowa State University Press; 1974.
6. Heimendinger J, Thompson B, Ockene J, et al. Reducing the risk of cancer through worksite intervention. *State Art Rev Occup Med.* 1990;5:707–723.
7. Block G, Lanza E. Dietary fiber sources in the United States by demographic group. *J Natl Cancer Inst.* 1987;79:83–91.

APPENDIX—Treatwell Eating Pattern Guidelines

TRIM FAT

Choose fish, skinned turkey and chicken, and trimmed lean red meat. Keep the amount at six ounces or less, cooked, a day.
Choose low-fat dairy products: skim, 1%, or 2% milk; low-fat yogurt; ice milk; and low-fat cheeses.
Use half the amount of fat or oil you normally use in cooking and baking, and at the table.

ADD FIBER

Eat at least one serving of high-fiber cereal every day.
Eat one or more servings of fruit at each meal and for snacks.
Eat one or more servings of vegetable at lunch and dinner.
Eat at least one of these foods at each meal:
Whole-grain bread, rice and/or pasta, potato
Dried beans, peas, or lentils

Health Status of Older Male Prisoners: A Comprehensive Survey

Patricia L. Colsher, PhD, Robert B. Wallace, MD, Paul L. Loeffelholz, MD, and Marilyn Sales, RN

ABSTRACT

Although prison populations remain predominantly young, the number of older inmates in state and federal correctional facilities has been increasing. This report describes the results of a health survey of 119 male inmates 50 years of age and older residing in Iowa state correctional facilities. Subjects' disease history included hypertension (40%), myocardial infarction (19%), and emphysema (18%). Most participants (97%) had missing teeth, 42% had gross physical functional impairments, and 70% smoked cigarettes. These findings have implications for health care provision and release planning. (*Am J Public Health.* 1992;82:881–884)

Introduction

The number of older inmates in state and federal correctional facilities has been increasing as a result of the aging of the general population and changes in arrest and sentencing patterns.[1–5] Although older persons still account for less than 4% of the US inmate population, there is concern about their impact on corrections systems. It has been suggested that older inmates may require additional health care services; assistance with personal care; special dietary, educational, rehabilitative, and recreational services; or modifications in the physical environment.[1,6–9] Concerns have also been raised about the need for special services upon parole or discharge.[9–11] However, because relatively little is known about the health and functional status of older inmates,[1] planning for this population is difficult. This report describes a population-based health survey of older inmates in Iowa correctional facilities.

Methods

Respondents

The respondents were 119 men (81% of the target population) aged 50 years and older who were residing in seven Iowa state correctional facilities housing men between June and December, 1989. Nonparticipation was most often the result of refusal by the inmate (23 of 29 nonparticipants). There were too few women residing in the one women's facility (n = 10) to tabulate here. The demographic characteristics of participants are shown in Table 1.

Interview Procedures and Measures

Letters describing the study were distributed to inmates, who returned postcards to investigators if they wanted to participate. Inmates were interviewed in person at the prisons; interviewers used a standardized questionnaire that required 90 minutes to administer. Interview contents are summarized in Appendixes 1 and 2.

Corrections staff were nearby but did not participate. Each participant received $5.00 compensation. The study was approved by the University of Iowa's institutional review board.

Statistical Analyses

Initial analyses[24] examined the distributions of measures according to age group (50 to 59 years, 60 years and older), using analysis of variance and chi-square for tests of statistical significance, as appropriate. Measures of cognitive function were adjusted for educational attainment.

Results

Table 2 shows the age group–specific and overall percentages of specific illnesses reported. The most common chronic illnesses were arthritis, hypertension, ulcers, prostate problems, and myocardial infarction. Rates were typically higher among inmates aged 60 years and older than among those aged 50 to 59 years.

Measures of age group–specific and overall health and functional status are presented in Table 3. Although almost 65% rated their health as excellent or good, almost half reported that their health had worsened since incarceration. Chest pain was reported by 61%, but fewer than 8% had a history of angina.[15] Fewer than 3% had all their teeth. Although 42% reported limitations in gross physical functional status, only 11% had limitations in routine self-care activities. Examination of the age group–specific measures indicates that participants aged 60 years and older were more likely to experience incontinence ($P < .01$), report hearing problems ($P < .05$), and have difficulty

Patricia L. Colsher and Robert B. Wallace are with the Department of Preventive Medicine and Environmental Health at the University of Iowa in Iowa City; Paul L. Loeffelholz and Marilyn Sales are with the Iowa Medical and Classification Center, Iowa Department of Corrections, in Oakdale.

Requests for reprints should be sent to Patricia L. Colsher, PhD, Department of Preventive Medicine, S404 Westlawn, University of Iowa, Iowa City, IA 52242.

This paper was submitted to the Journal July 23, 1990, and accepted with revisions November 26, 1991.

TABLE 1—Demographic Characteristics (n = 119)

Age, y	
Mean	57.0 (SD = 6.3)
Median	56.0
Educational attainment, y	
Mean	10.4 (SD = 3.4)
Median	10.0
Marital status, n	
Divorced/legally separated	54 (47.4%)
Married	36 (31.0%)
Never married	16 (13.8%)
Widowed	9 (7.8%)
Race, n	
White	99 (83.1%)
Black	11 (9.6%)
Native American	2 (1.7%)
Hispanic	2 (1.7%)
Prior prison sentence, n	69 (62.2%)
Life sentence, n	21 (18.1%)
Time already served, y	
Mean	5.7 (SD = 6.7)
Median	3.0
Time life to serve, y	
Mean	11.9 (SD = 11.4)
Median	9.0

Note. Information on marital status was missing for 4 persons, on race for 5 persons, on prior sentence for 8 persons, and on current sentence for 3 persons.

TABLE 2—Percentage of Male Inmates with Lifetime History of Specific Self-reported Physician-diagnosed Illness

	Age, y		
	50–59 (n = 82)	> 59 (n = 37)	Overall (n = 119)
Arthritis	40.2	56.8	45.4
Hypertension	36.7	45.9	39.7
Any venereal disease	21.5	21.6	21.6
Stomach or intestinal ulcers	18.3	27.0	21.0
Prostate problems	17.1	27.0	20.2
Myocardial infarction	17.7	21.6	19.0
Emphysema	14.6	27.0	18.5
Diabetes	10.1	13.5	11.2
Asthma	8.5	10.8	9.2
Stroke	3.8	16.2	7.8
Cancer	6.3	8.1	6.9
Cirrhosis or liver disease	4.9	2.7	4.2
Injury requiring medical care	78.5	73.0	76.7

TABLE 3—Health and Functional Status According to Age Group

	Age, y		
	50–59 (n = 82)	> 59 (n = 37)	Overall (n = 119)
Mean no. major illnesses (of 5)	0.7 (.10)[a]	1.1 (.14)	0.8 (.08)
% with at least one	59.8	73.0	63.9
Weight, lb	189.2 (3.6)	176.5 (5.4)	185.2 (3.1)
Body mass index	3.86 (.06)	3.68 (.10)	3.81 (.06)
Self-perceived health status			
% Excellent or good	69.5	54.1	64.7
Health since incarceration			
% Better	3.8	8.1	5.2
% Worse	44.3	54.1	47.4
Mean no. respiratory symptoms (of 9)	2.5 (.29)	3.3 (.42)	2.8 (.23)
% Chest pain	57.0	70.3	61.2
% Angina[12]	7.3	8.1	7.6
% Low back pain in past year	27.8	35.1	30.2
% Any incontinence	13.9	37.8	21.6
% Missing all teeth	34.2	48.6	38.7
% Missing some teeth	63.3	48.7	58.7
% Recognize friend across street	94.9	81.2	93.1
% Read normal newsprint	96.2	86.5	93.1
% Have hearing aid	6.3	16.2	9.5
% Believes have hearing loss	30.8	55.6	38.6
% Trouble hearing normal voices	9.1	33.3	16.8
% With gross physical functional disability	36.0	51.4	42.2
% With routine self-care dependency	8.9	16.2	11.2
% With dependency in instrumental activities	9.0	24.3	20.7
% With difficulty making isolated movements	68.3	83.8	73.2

[a]Values in parentheses are standard errors of the mean.

hearing normal voices ($P < .01$), than were those aged 50 to 59 years.

Psychobehavioral and social measures are summarized in Table 4. Participants typically rated their own memory and cognitive function as excellent or good, and of the 8 items on the mental status screening examination, they had a mean of 7.0 correct. Although older men were somewhat less likely than younger men to rate their own cognitive function as excellent or good ($P < .1$), there were no systematic differences in performance measures of cognitive function. Depressive and anxiety symptoms were common, but psychotic symptoms were less frequently reported. Older men reported significantly more offspring ($P < .05$), but there were no differences in visitation patterns. Only 37% of the men had visitors during the month prior to the interview, and most wished they had more.

Most inmates (69.8%) were current cigarette smokers, and 18.5% were former smokers. Current smokers consumed a mean of 23.0 cigarettes (SD = 13.6) per day. Almost all inmates (96.6%) had a history of alcohol consumption, and 45.4% described themselves as having been heavy drinkers; 29.4% reported illicit drug use.

Discussion

This report describes the health and functional status of older inmates in Iowa state correctional facilities. Although chronic illness was common and many inmates reported limitations in gross physical functional ability, most inmates were able to perform routine self-care activities and described their own health as excellent or good. There was little evidence of cognitive impairment. Sizable social networks were described, but visits were not common.

Self-reported rates of lifetime substance use were high and may have implications for health care. The high rates of smoking (over twice the national average)[25] may contribute to the respiratory symptoms and illnesses, cardiovascular

conditions, and cancer found among this population; thus, smoking cessation programs appropriate to the prison environment should be developed. A history of heavy drinking was four times as common among these inmates as it was in a population-based study of older Iowans.[26] In that study, a history of heavy drinking was strongly associated with increased morbidity and mortality. The effects of illicit drug use on older persons remains largely uninvestigated. However, the absence of substantial alcohol and other substance abuse in the prison system could have a mitigating effect on subsequent disease occurrence.

The finding of increased rates of incontinence, sensory impairment, and flexibility impairment suggests that corrections departments may need to modify the physical environment of prisons that house large numbers of older inmates and provide special programming.[1,13] The provision of specialized equipment, however, may raise issues of limited resources and priority setting.

Several limitations should be noted in these findings. First, the cohort is small although it does include most older male inmates in Iowa state correctional facilities. Second, some misclassification may have been introduced by the use of self-reported history of chronic conditions. Third, Iowa has lower crime and incarceration rates than do many other states,[27] and the prisons in Iowa are relatively small. Moreover, inmates in Iowa facilities may have different levels of access to or use of health care services than inmates in other states. Thus, the findings reported here should be replicated in other prison systems, especially those with larger numbers of minority inmates and women. □

TABLE 4—Psychobehavioral and Social Measures According to Age Group

	Age, y		
	50–59 (n = 82)	> 59 (n = 37)	Overall (n = 119)
Self-rated memory			
% Excellent or good	73.6	74.6	74.1
Self-rated cognition			
% Excellent or good	89.5	77.3	85.8
Recall memory test, no. items (of 20)	5.8 (.27)[a]	5.9 (.41)	5.8 (.23)
Recognition memory test, no. items (of 20)	10.6 (.49)	10.2 (.74)	10.5 (.41)
Mental status exam, no. correct (of 8)	7.1 (.12)	6.9 (.17)	7.0 (.10)
Depressive symptoms (range: 11–33)	15.0 (.44)	16.3 (.64)	15.4 (.36)
Loneliness symptoms (range: 4–12)	7.6 (.14)	7.9 (.22)	7.1 (.12)
Anxiety symptoms (range: 5–15)	8.2 (.26)	8.1 (.40)	8.1 (.22)
% Worry "too much"	14.6	18.9	16.0
Psychotic symptoms (range: 0–12)	1.3 (.19)	1.4 (.28)	1.3 (.16)
Mean no. offspring	2.9 (.37)	4.5 (.54)	3.4 (.31)
Mean no. close relatives	6.0 (1.2)	3.9 (1.8)	5.3 (1.0)
Mean no. friends in prison	5.1 (1.3)	8.8 (1.9)	6.3 (1.1)
Mean no. friends outside prison	8.1 (1.9)	11.6 (2.7)	9.3 (1.5)
Mean no. visitors in past month	1.2 (.23)	1.3 (.42)	1.2 (.23)
% With any visitor	39.2	32.4	37.1
% Wish had more visitors	67.1	59.4	64.7
% Prefer solitary activities	29.3	27.0	28.6
% Talk about problems	32.9	43.2	36.1

Note. Cognitive function measures are adjusted for educational attainment.
[a]Values in parentheses are standard errors of the mean.

Acknowledgments

This research was supported by National Institute on Aging grant AG-07094, the University of Iowa Teaching Nursing Home.

Portions of this research were presented at the 1989 meeting of the Gerontological Society of America and the 1990 meeting of the American Geriatrics Society.

The authors gratefully acknowledge the contributions and assistance of staff and inmates at Iowa Department of Corrections facilities: Clarinda Correctional Facility (Mark Lund, warden; Ann Taylor, contact person); Iowa Correctional Institute for Women (Barbara Olk, superintendent; Cheryl Overlin, contact person); Iowa Medical and Classification Center (Rusty Rogerson, warden; Lowell Brandt and Kay Rhoads, contact persons); Iowa Men's Reformatory (John Thalacker, warden; Jen Kula, contact person); Iowa State Penitentiary (Crispus Nix, warden; Ron Welder, assistant deputy warden; Scott Stevens, contact person); Mt. Pleasant Correctional Facility (David Scurr, superintendent; Sue Erue, contact person); North Central Correctional Facility (John Ault, superintendent; Susan Meyer, contact person); Riverview Release Center (John Mathis, superintendent; Maria Adams, contact person). Thanks are also extended to Edward O'Brien for his assistance and advice.

References

1. Dugger RL. The graying of America's prisons. *Corrections Today*. 1988;(June):26, 28–31,34.
2. Feinberg G. White-haired offenders: an emergent social problem. In: Willbanks W, Kim PHK, eds. *Elderly Criminals*. Lanham, Md: University Press of America; 1984:83–108.
3. Newman DJ. Elderly offenders and American crime patterns. In: Newman ES, Newman DJ, Gerwitz ML, eds. *Elderly Criminals*. Cambridge, Mass: Oelgeschlager, Gunn, and Hain; 1984:3–16.
4. Sapp AD. Arrests for major crimes: trends and patterns for elderly offenders. In: Chaneles S, Burnett C, eds. *Older Offenders: Current Trends*. New York, NY: Haworth Press; 1989:19–44.
5. Shichor D. The extent and nature of lawbreaking by the elderly: a review of arrest statistics. In: Newman ES, Newman DJ, Gerwitz ML, eds. *Elderly Criminals*. Cambridge, Mass: Oelgeschlager, Gunn, and Hain; 1984:17–32.
6. Goetting A. Prison programs and facilities for elderly inmates. In: Newman ES, Newman DJ, Gerwitz ML, eds. *Elderly Criminals*. Cambridge, Mass: Oelgeschlager, Gunn, and Hain; 1984:169–175.
7. Golden D. Elderly offenders in jail. In: Newman ES, Newman DJ, Gerwitz ML, eds. *Elderly Criminals*. Cambridge, Mass: Oelgeschlager, Gunn, and Hain; 1984:143–152.
8. Moore EO. Prison environments and their impact on older citizens. In: Chaneles S, Burnett C, eds. *Older Offenders: Current Trends*. New York, NY: Haworth; 1989: 175–192.
9. Rubenstein D. The elderly in prison: a review of the literature. In: Newman ES, Newman DJ, Gerwitz ML, eds. *Elderly Criminals*. Cambridge, Mass: Oelgeschlager, Gunn, and Hain; 1984:153–168.
10. Goetting A. The elderly in prison: issues and perspectives. *J Res Crime Delinquency*. 1983;(July):291–309.
11. Booth DE. Health status of the incarcerated elderly: issues and concerns. In: Chaneles S, Burnett C, eds. *Older Offenders: Current Trends*. New York, NY: Haworth; 1989:193–214.
12. American Thoracic Society Committee on Standards for Epidemiologic Surveys in Chronic Respiratory Disease. *Standards for Epidemiologic Surveys in Chronic Respiratory Disease*. New York, NY: National Tuberculosis and Respiratory Disease Association; 1978.
13. US Bureau of the Census. *Health Interview Survey*. Washington, DC: US Govt Printing Office; 1979.
14. Cornoni-Huntley J, Brock DB, Ostfeld AM, Taylor JO, Wallace RB. *Established Populations for Epidemiologic Studies of the Elderly: Resource Data Book*. Bethesda, Md: National Institutes of Health; 1986.
15. Rose GA, Blackburn H. *Cardiovascular*

Survey Methods. Geneva, Switzerland: World Health Organization; 1968.
16. US Department of Health, Education, and Welfare. HANES I augmentation survey of adults 25–74 years, United States, 1974–1975. *Vital and Health Statistics.* Series 1, No. 14. Washington, DC: US Government Printing Office; 1978. DHEW publication (PHS) 78–1314.
17. Rosow I, Breslau N. A Guttman health scale for the aged. *J Gerontology.* 1966;21: 556–559.
18. Katz S, Apkom C. A measure of primary sociobiological function. *Int J Health Serv.* 1976;6:493–508.
19. Nagi SZ. An epidemiology of disability among adults in the United States. *Milbank Mem Fund Q.* 1976;54:439–468.
20. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *J Am Psychiatr Soc.* 1975;23:433–441.
21. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1: 385–401.
22. Russell D, Paplau L, Cutrona C. The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence. *J Pers Soc Psychol.* 1980;39:472–480.
23. Blazer D, George LK, Hughes D. Schizophrenic symptoms in an elderly community population. In: Brody JA, Maddox GL, eds. *Epidemiology and Aging: An International Perspective.* New York, NY: Springer; 1988:134–149.
24. Statistical Analysis Systems. *SAS Guide.* Cary, NC: Statistical Analysis Systems; 1985.
25. Thornberry OT, Wilson RW, Golden PM. Health promotion data for the 1990 objectives. Advance data from *Vital and Health Statistics,* No. 126. Hyattsville, Md: National Center for Health Statistics; 1986.
26. Colsher PL, Wallace RB. Correlates and consequences of history of heavy alcohol consumption among elderly men. *J Stud Alcohol.* 1990;51:528–535.
27. US Bureau of the Census. *Statistical Abstract of the United States: 1989.* Washington, DC: US Government Printing Office; 1989.

APPENDIX 1—Summary of Demographic, Health, and Functional Status Measures

Demographic characteristics
- Age
- Educational attainment
- Race
- Marital status
- Criminal history

General health status
- Self-reported lifetime history of physician diagnosed illness
  - myocardial infarction, stroke, hypertension, cancer, diabetes, arthritis, asthma, emphysema, ulcers, cirrhosis, venereal disease, prostate problems
- Symptoms and complaints in past year
  - dry or itching skin, pain in legs at night, indigestion, fallen arches, bunions or corns on feet, other foot problems, trouble with fingernails or toenails, joint pain, morning joint stiffness
- Enumeration of major illnesses
  - myocardial infarction, stroke, hypertension, cancer, diabetes
- Respiratory symptoms[12]
  - shortness of breath, coughing, phlegm production
- Self-perceived health status
  - as compared with age mates,[13] as changed since incarceration
- Sleep problems[14]
  - trouble falling asleep, wake during night, wake too early, feel rested in morning, need to nap during day
- London School of Hygiene Cardiovascular Questionnaire[15]
- Urinary incontinence[14]
- Low back pain in past year[14]
- Self-rated vision function[16]
- Self-rated hearing function[16]
- Self-reported height and weight
- Tooth loss[14]

Physical functional status
- Gross physical function[17]
  - walk a half-mile, walk up stairs, do heavy work
- Instrumental activities
  - fill out forms and requests, handle own canteen account, clean cell, make telephone calls
- Routine self-care activities[18]
  - bathe, dress, groom, toilet, eat, walk across small room, transfer from bed to chair
- Isolated movements[19]
  - push heavy objects, stoop, lift 10 pounds, reach arms above shoulders, write, handle small objects, stand in place for 15 minutes, get up from chair, stand on one foot, pick up object from floor

APPENDIX 2—Summary of Substance Use and of Psychobehavioral and Social Measures

Substance use[16]
- Tobacco
  - cigarettes, cigars, pipes, smokeless tobacco
- Alcohol[14]
  - whether alcohol ever consumed, history of heavy drinking
- Illicit drugs
  - stimulants, sedatives, opiates, marijuana, hallucinogens

Psychobehavioral function
- Cognitive function
  - mental status screening examination,[20] 20-item memory performance test,[14] self-rated memory,[14] self-rated cognition
- Affect and psychiatric symptoms
  - Center for Epidemiologic Studies Depression Scale,[21] anxiety symptoms,[14] loneliness,[22] psychotic symptoms[23]

Social network/style[14]
- Number of offspring
- Number of close relatives
- Number of friends inside prison
- Number of friends outside prison
- Number of visitors in past month
- Desire for more visitors
- Tendency to discuss problems with others
- Preference for solitary activities

A Neonatal Hepatitis B Surveillance and Vaccination Program: New York City, 1987 to 1988

Kelly J. Henning, MD, Daphna M. Pollack, MPH, and Stephen M. Friedman, MD, MPH

ABSTRACT

From July 1987 to June 1988, 1030 pregnant women with hepatitis B were reported to a New York City surveillance program. Among 832 infants under follow-up, the coverage rates for combined hepatitis B immune globulin and vaccine doses 1, 2, and 3 were 84%, 77%, and 59%, respectively. Infants covered by Medicaid and uninsured Black and Hispanic infants were significantly less likely to be completely vaccinated. An estimated 160 cases of chronic hepatitis B infection were prevented among infants enrolled in the program. Strategies are needed to improve vaccine coverage among hard-to-reach groups. (*Am J Public Health.* 1992;82:885–888)

Introduction

Hepatitis B virus (HBV) transmission from mother to infant during the perinatal period is highly efficient.[1–4] If the infant is not infected in the postnatal period, subsequent exposure to the mother or to other HBsAg-positive family members can lead to infection and chronic carriage.[5,6] Treatment of infants born to HBsAg-positive mothers with hepatitis B immune globulin (HBIG) and hepatitis B (HB) vaccine results in an 85% to 95% reduction in the carrier state.[7–9]

Women at high risk for hepatitis B carriage include women of Asian or Pacific Island descent, women born in Haiti or sub-Saharan Africa, and women who are intravenous drug abusers. In 1986, almost 2500 infants were born to high-risk mothers in New York City.[10] In the same year, the New York City Department of Health (NYCDOH) began a surveillance and vaccination program for HBsAg-positive pregnant women and their newborns.

Methods

The surveillance system for HBsAg-positive pregnant women included both active and passive components. Active surveillance targeted 13 New York City hospitals (8 municipal and 5 voluntary) serving high-risk populations. Yearly, about 40 000 births, representing 30% of all births in the city, occurred at these 13 hospitals. Five of the hospitals served large Asian communities. NYCDOH field workers reviewed hospital laboratory logs monthly and collected identifying information for HBsAg-positive pregnant women. Passive reporting sources included Maternity Infant Care, a New York State and federally funded family planning project; a large health maintenance organization with 895 000 subscribers; private physicians; and nonprogram hospitals.

NYCDOH workers contacted the HBsAg-positive women by telephone, letter, or home visit. A questionnaire was administered and women were counseled (in English, Spanish, Chinese, French, or Creole) regarding HBV transmission, the importance of screening household contacts, and the need for neonatal HB vaccination.

To facilitate administration of HBIG and the first dose of HB vaccine, all women enrolled in the program were given a letter from the NYCDOH stating their surface antigen status. Women were encouraged to take this letter to the hospital at the time of delivery. An infant was considered to have been given HBIG and vaccine dose 1 if a form documenting administration was received by the program or if the mother could clearly describe the administration of vaccine. Each infant was then followed, via multiple telephone calls to the mother and to the infant's caregiver, to ensure vaccination at 1 and 6 months of age. Infants for whom receipt of vaccine dose 3 could not be documented by 18 months of age were considered lost to follow-up.

The surveillance program was budgeted at less than $200 000 per year; screening of the pregnant women, provision of HBIG, and provision of the first dose of HB vaccine were not covered by the program.

Statistical analysis was done using standard microcomputer software (EpiInfo, Centers for Disease Control) and exact confidence intervals for relative risks were calculated using the methods of Thomas and Gart, where appropriate.[11]

The authors are with the New York City Department of Health, Bureau of Immunization, Epidemiology and Prevention Services. Kelly J. Henning is also with the US Centers for Disease Control, Division of Field Epidemiology, Epidemiology Program Office, Atlanta, Ga.

Requests for reprints should be sent to Stephen M. Friedman, MD, MPH, New York City Department of Health, 125 Worth Street, Box 21, New York, NY 10013.

This paper was submitted to the Journal December 17, 1990, and accepted with revisions November 27, 1991.

TABLE 1—HBIG and Vaccine Coverage by Dose for 832 Infants, Hepatitis B Surveillance and Vaccination Program, New York City, July 1987 to June 1988

HBIG/Vaccine Dose	Infants Receiving HBIG/Vaccine	
	No.	%
HBIG	706	85
1	810	97
2	738	89
3	554	67
HBIG + 1	702	84
HBIG + 1 + 2	637	77
HBIG + 1 + 2 + 3	492	59

Results

In the 1-year period from July 1, 1987, through June 30, 1988, 1030 HBsAg-positive women were identified. The 13 participating hospitals screened 20 108 pregnant women; 488 (2.4%) were HBsAg positive. Three hundred forty-two HBsAg-positive women were reported from passive surveillance sources. The remaining 200 HBsAg-positive women were identified in the postpartum period and reported to the NYCDOH shortly after delivery (median, 10 days).

The study population consisted of 830 (81%) of the 1030 positive women. Reasons for exclusion from further study were as follows: 89 women did not respond to health department calls or letters, 50 provided insufficient locating information, 30 moved before delivery, 26 seroconverted to antigen-negative status before delivery, and 3 aborted or miscarried.

The 830 women delivered 832 infants. The combined HBIG and vaccine dose 1, 2, and 3 coverage rates were 84%, 77%, and 59%, respectively (Table 1). Delivery of HBIG and three doses of vaccine was similar among infants of women identified during the antepartum and postpartum periods (61% vs 56%, $P = .18$). The majority (80%) of infants received HBIG and the subsequent vaccine doses within the recommended dosage schedule.[12]

Questionnaires were completed on 670 (81%) of the 830 HBsAg-positive women. Vaccine completion varied by maternal race/ethnicity, maternal birthplace, and health insurance (Table 2). Infants of Black and Hispanic women were 24% and 37% less likely, respectively, to complete the series than were Asian infants. Infants of women who were Medicaid recipients were almost 30% less likely to complete the series than were infants covered by private insurance.

To evaluate health insurance as a possible confounder for the racial differences observed, vaccine completion rates by maternal race were stratified by health insurance (Table 3). Vaccine completion by race among women with private insurance or Medicaid was similar; however, in

TABLE 2—Vaccine Completion Rates by Maternal Characteristics, Hepatitis B Surveillance and Vaccination Program, New York City, July 1987 to June 1988

Maternal Characteristic	Total Infants[a]	Vaccine Series Completed		RR	95% CI	*P* Value
		No.	%			
Age, y						
<18	63	38	60	0.90	0.73, 1.11	.31
≥18	607	405	67	1.00	reference	...
Race/ethnicity						
Asian	411	304	74	1.00	reference	...
Black	158	89	56	0.76	0.66, 0.88	<.005
Hispanic	73	34	47	0.63	0.49, 0.81	<.005
White	23	14	61	0.82	0.59, 1.15	.17
Maternal birthplace						
Asia[b]	355	265	75	1.00	reference	...
United States	70	35	50	0.67	0.53, 0.85	<.005
Haiti	56	38	68	0.91	0.75, 1.10	.28
Dominican Republic	39	17	44	0.58	0.41, 0.84	<.005
Pacific Island	13	10	77	1.03	0.62, 1.27	1.00[d]
Other[c]	80	46	58	0.77	0.63, 0.94	<.005
Health insurance						
Private	247	182	74	1.00	reference	...
None	174	123	71	0.96	0.85, 1.08	.50
Medicaid	123	66	54	0.73	0.61, 0.87	<.005
Intravenous drug abuse						
Yes	11	6	55	0.80	0.36, 1.20	.34[d]
No	568	388	69	1.00	reference	...
Total	630	443	66			

Note. RR = relative risk; CI = confidence interval.
[a]Groups by factors may not add to 670 because of nonresponses.
[b]Asian birthplaces include Burma, Cambodia, China, Hong Kong, India, Japan, Korea, Laos, Pakistan, Singapore, Taiwan, Thailand, and Vietnam.
[c]Other birthplaces include Antigua, Afghanistan, Barbados, Belize, Colombia, Costa Rica, Dominica, Ecuador, Grenada, Greece, Guyana, Honduras, Israel, Italy, Jamaica, Libya, Niger, Nigeria, Nauru, Panama, Poland, Puerto Rico, Saudi Arabia, Sierra Leone, Trinidad, USSR, Venezuela, and Yugoslavia.
[d]Exact CI and two-tailed Fisher's exact p value.

TABLE 3—Vaccine Completion Rates by Maternal Race, Stratified by Health Insurance, Hepatitis B Surveillance and Vaccination Program, New York City, July 1987 to June 1988

Maternal Race/Ethnicity	Total Infants	Vaccine Series Completed		RR	95% CI	*P* Value
		No.	%			
Private Health Insurance						
Asian	194	147	76	1.00	reference	...
Black	33	21	64	0.84	0.64, 1.10	0.14
Hispanic	17	12	71	0.93	0.68, 1.28	0.64
White	3	2	67	0.88	0.16, 1.30	0.57*
No Health Insurance						
Asian	111	89	80	1.00	reference	...
Black	42	23	55	0.68	0.51, 0.91	<0.005
Hispanic	13	5	38	0.48	0.24, 0.96	<0.005
White	8	6	75	0.94	0.45, 1.23	0.66*
Medicaid						
Asian	31	18	58	1.00	reference	...
Black	57	31	54	0.94	0.64, 1.37	0.74
Hispanic	27	12	44	0.77	0.46, 1.28	0.30
White	7	4	57	0.98	0.32, 1.75	1.00*

Note. RR = relative risk; CI = confidence interval.
*Exact CI and two-tailed Fisher's exact *P* value.

the uninsured category, infants of Black and Hispanic women remained much less likely to complete the vaccine series than were Asian infants.

To evaluate the impact of this program, we compared completion rates for infants who received the first vaccine dose at the 13 program hospitals with those who received the first dose at nonprogram hospitals. Although 373 (73%) of the 514 infants who received dose 1 at program hospitals went on to complete the series, only 68 (59%) of the 116 infants who received dose 1 at nonprogram hospitals did so (relative risk = 1.24, 95% confidence interval = 1.05, 1.46). This difference persisted after we adjusted for maternal race and health insurance coverage.

We used the following assumptions to estimate the number of cases of chronic HBV infection prevented: (1) a 25% HBeAg positivity rate for the HBsAg-positive women[2,13]; (2) in the absence of vaccination, a 90% HBV transmission rate to infants of HBeAg-positive women, with a 90% chronic carriage rate for these infants[1,2]; (3) in the absence of vaccination, a 20% chronic carriage rate for infants born to HBeAg-negative women; and (4) a vaccine efficacy of 92%.[3,4,13] Among the 493 infants who completed the vaccine series, we estimate that 174 cases of chronic HBV infection would have occurred had they not been vaccinated. We expect that vaccination prevented 160 (92%) cases.

Discussion

This is the first report of a large surveillance system and vaccination program for hepatitis B–infected pregnant women and their newborns in the continental United States.

A primary aim of the program is to assist in delivery of the second and third dose of HB vaccine. Although the second dose of vaccine was administered to 89% of the infants enrolled in the program, only 67% received the third vaccine dose. A similar decline was noted among Asian and Pacific Island women who received county medical services in California.[13] In the New York City program, as in the California study, most infants without a documented third dose of vaccine were lost to follow-up. Because some of these infants may have completed the vaccine series elsewhere, the overall completion rate of 59% is likely to be an underestimate. However, many infants lost to follow-up had been receiving care at public clinics and hospitals and would have been unlikely to complete the series at private sites.

Method of health care payment explained most of the differences in vaccine completion rates. Medicaid-insured infants and uninsured Black and Hispanic infants had very low vaccine completion rates. Medicaid reimbursement in New York State falls 60 cents short of vaccine cost, does not adequately cover the cost of vaccine administration, and requires months for release of funds—all barriers to physicians caring for infants requiring HB vaccination. Improvement of vaccine completion will require increased Medicaid reimbursement.

The racial differences noted among uninsured mothers probably reflect socioeconomic differences. More than 95% of the uninsured Asian and White mothers gave birth at private hospitals; these women may be in a higher socioeconomic group than the uninsured Black and Hispanic mothers, most of whom delivered in public hospitals. Intensive follow-up and ready access to free vaccine may improve coverage rates for infants born in public hospitals. Additional strategies to improve completion rates include providing documentation of HB vaccine on the immunization card and incorporating HB vaccine into routine childhood vaccination schedules. HB vaccine coverage rates as high as 75% to 95% have been achieved when HB vaccine is administered with other childhood vaccines.[14–17]

In 1988, the Immunization Practices Advisory Committee recommended universal screening of all pregnant women for HBsAg.[12] In response, New York State passed a law, which took effect in May 1990, mandating universal screening of all pregnant women.[18] Future directions for the program include (1) administration of the second and third dose of HB vaccine at 2 and 6 months of age, concomitant with the diphtheria-tetanus-pertussis vaccine, for those vaccinated at city clinics; (2) HB surface antibody testing at 12 months; and

(3) provision of HBsAg screening and HB vaccine to family contacts of HBsAg-positive pregnant women.

The NYCDOH program was designed to establish surveillance in hospitals that serve women at highest risk for hepatitis B carriage. Therefore, the findings reported here do not reflect all HBsAg carrier mothers in New York City. However, we believe these findings may apply to other large, diverse cities in the United States. Health departments can assist physicians and patients by coordinating the transfer of information between physician groups, providing education to HBsAg-positive women, and ensuring follow-up of infants, particularly those in hard-to-reach groups. □

Acknowledgments

Preliminary results of this study were presented at the 39th Annual Conference of the Epidemic Intelligence Service, Atlanta, Ga, April 23–27, 1990.

The authors wish to acknowledge Dr Ismat Aziz for her work in the early stages of the program and Drs Robert Gunn and Harold Margolis for their comments on the manuscript. We also thank Kim Mealing for her technical assistance.

References

1. Stevens CE, Beasley RP, Tsui J, Lee W-C. Vertical transmission of hepatitis B antigen in Taiwan. *N Engl J Med.* 1975;292:771–774.
2. Stevens CE, Toy PT, Tong MJ, et al. Perinatal hepatitis B virus transmission in the United States: prevention by passive-active immunization. *JAMA.* 1985;253:1740–1745.
3. Xu ZY, Liu CB, Francis DP, et al. Prevention of perinatal acquisition of hepatitis B virus carriage using vaccine: preliminary report of a randomized, double-blind placebo-controlled and comparative trial. *Pediatrics.* 1985;76:713–718.
4. Beasley RP, Trepo C, Stevens CE, Szmuness W. The e antigen and vertical transmission of hepatitis B surface antigen. *Am J Epidemiol.* 1977;105:94–98.
5. Kachiwagi S, Hayashi J, Ikematsu H, et al. Transmission of hepatitis B virus among siblings. *Am J Epidemiol.* 1984;120:617–625.
6. Franks AL, Berg CJ, Kane MA, et al. Hepatitis B virus infection among children born in the United States to southeast Asian refugees. *N Engl J Med.* 1989;321:1301–1305.
7. Beasley RP, Hwang L-Y, Lee GC-Y, et al. Prevention of perinatally transmitted hepatitis B virus infections with hepatitis B immune globulin and hepatitis B vaccine. *Lancet.* 1983;2:1099–1102.
8. Wong VCW, Ip HMH, Reesink HW, et al. Prevention of the HBsAg carrier state in newborn infants of mothers who are chronic carriers of HBsAg and HBeAg by administration of hepatitis-B vaccine and hepatitis-B immunoglobulin: double-blind randomised placebo-controlled study. *Lancet.* 1984;1:921–926.
9. Stevens CE, Taylor PE, Tong MJ, et al. Yeast-recombinant hepatitis B vaccine: efficacy with hepatitis B immune globulin in prevention of perinatal hepatitis B virus transmission. *JAMA.* 1987;257:2612–2616.
10. Friedman SM, DeSilva LP, Fox HE, Bernard G. Hepatitis B screening in a New York City obstetrics service. *Am J Public Health.* 1988;78:308–310.
11. Thomas DG, Gart JJ. A table of exact confidence limits for differences and ratios of two proportions and their odds ratios. *J Am Stat Assoc.* 1977;72:73–76.
12. Centers for Disease Control, Immunization Practices Advisory Committee. Prevention of perinatal transmission of hepatitis B surface antigen. *MMWR.* 1988;37:341–346.
13. Klontz KC. A program to provide hepatitis B immunoprophylaxis to infants born to HBsAg-positive Asian and Pacific Island women. *West J Med.* 1987;146:195–199.
14. The Gambia Hepatitis Study Group. Hepatitis B vaccine in the expanded programme of immunisation: the Gambian experience. *Lancet.* 1989;1:1057–1060.
15. Goh KT, Doraisingham S, Tan KL, et al. The hepatitis B immunization programme in Singapore. *Bull WHO.* 1989;67:65–70.
16. Schalm SW, Mazel JA, deGast GC, et al. Prevention of hepatitis B infection in newborns through mass screening and delayed vaccination of all infants of mothers with hepatitis B surface antigen. *Pediatrics.* 1989;83:1041–1048.
17. Waters JR. Universal prenatal screening for hepatitis B, Alberta, 1985–1988. *Can Disease Weekly Rep.* 1989;15:29–32.
18. 1990 NY Laws ch 4, §2500-e.

Physician and Infection Control Practitioner HIV/AIDS Reporting Characteristics

ABSTRACT

We surveyed a random sample of South Carolina physicians and infection control practitioners about the reporting of human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS) cases. Of physicians surveyed, 79% indicated that HIV infection as well as AIDS should be reported by name. The following characteristics were associated with those physicians who do not report AIDS cases: not feeling responsible for reporting, not reporting a case perceived to have been reported in another state, believing that information required for reporting is not on the chart, and residing in an urban setting. Targeted education can address these underreporting factors. (*Am J Public Health.* 1992; 82:889–891)

Jeffrey L. Jones, MD, MPH, Pamela Meyer, MSPH, Carol Garrison, PhD, Lynda Kettinger, MPH, and Patricia Hermann, MD, MPH

Introduction

Reporting of acquired immunodeficiency syndrome (AIDS) is essential to understand disease transmission patterns and target preventive programs.[1–4] Reporting also affects state funding because disbursement of federal funds for several human immunodeficiency virus (HIV)/AIDS prevention and treatment programs is based on the annual number of cases reported.[5] To inquire about AIDS reporting practices, the South Carolina Department of Health and Environmental Control (DHEC) conducted a statewide survey of physicians and infection control practitioners (ICPs). Topics covered in the survey included the number of AIDS cases cared for, reporting attitudes and practices, and services DHEC could provide to facilitate reporting.

In South Carolina, AIDS has been reportable since 1982. HIV infection without AIDS has been reportable from physicians, hospitals, and laboratories since 1986. As of June 30, 1991, there were 1396 cases of AIDS and 5150 cases of HIV infection reported in South Carolina. Of the 1396 AIDS cases, 24 (1.7%) were in children younger than 13 years of age. With an annual AIDS incidence rate of 9.4 per 100 000 (June 1990 through May 1991), South Carolina ranked 23rd in the nation.[6]

Methods

In spring 1989, after implementation of active AIDS surveillance, the DHEC survey was mailed to (1) all physician specialists in the fields of infectious diseases, pediatric allergy/infectious diseases, hematology/oncology, and pulmonary medicine listed in the South Carolina Medical Association Directory; (2) an 18% random sample of the states' 1712 primary care internists, and family and general practitioners; and (3) ICPs from all inpatient facilities licensed by DHEC, military, and Veterans Administration hospitals. Physicians who identified themselves as ICPs were analyzed with ICPs (most of whom, in South Carolina, are nurses). Due to the small number (24) of reported pediatric AIDS cases, primary care pediatricians were not surveyed.

Three mailings of the questionnaire were made under different cover letters to enhance the response rate. There were 32 (11%) substitutions on the random sample for physicians who had retired or moved to another state since the list was compiled in the previous year. Chi-square for trend and prevalence ratios were calculated using Epi Info[7] software. SAS[8] software was used for backward stepwise logistic regression.

Results

Of the 542 questionnaires mailed, 439 (81%) were completed. The group response rates were physician specialists, 81% (114 of 140); primary care physicians, 78% (234 of 301); and ICPs, 90% (91 of 101). Of the 91 responding ICPs, 85 were nurses and 6 were physicians. There were no significant differences between responding and nonresponding physicians in years of experience, location of practice (rural or urban), or type of practice (specialist or primary care).

Physicians in larger hospitals were more likely than others to have cared for at least one person with AIDS (those in hospitals with more than 300 beds, 62%; with 101 to 300 beds, 51%; with 100 beds or fewer, 25%; $P < .001$ chi-square for linear trend). Forty-three percent of the surveyed primary care physicians, 72% of the physician specialists, and 64% of the ICPs

Jeffrey L. Jones, Pamela Meyer, Lynda Kettinger, and Patricia Hermann are with the South Carolina Department of Health and Environmental Control in Columbia, and Carol Garrison is with the University of South Carolina School of Public Health.

Requests for reprints should be sent to Jeffrey L. Jones, MD, MPH, State Epidemiologist, South Carolina Department of Health and Environmental Control, 2600 Bull Street, Columbia, SC 29201.

This paper was submitted to the Journal May 14, 1991, and accepted with revisions October 18, 1991.

TABLE 1—Percent of Surveyed Physicians, by Specialty, Reporting Having Cared for Specified Numbers of AIDS Patients

	Physician Specialty (n = 348)									
	Family Practice (126)	General Practice (32)	Internal Medicine (76)	Allergy & Immunology (5)	Pediatric Allergy (3)	Hematology (8)	Hematology/ Oncology (19)	Oncology (20)	Infectious Disease (22)	Pulmonary Disease (37)
No. AIDS patients care for										
0	63	69	43	100	34	50	21	40	9	22
1–4	34	22	34	0	66	25	32	30	13	19
5–9	2	3	13	0	0	13	37	25	14	13
10–14	1	3	6	0	0	0	5	5	18	22
15–19	0	0	1	0	0	12	5	0	0	0
20+	0	3	3	0	0	0	0	0	41	24

TABLE 2—Comparison of Physician Reporting Practices by Characteristics of Nonreporters[a]

Characteristics of Nonreporters	Report AIDS Cases Rarely or Never[b] No.	(%)	Report AIDS Cases Usually or Always[c] No.	(%)	PR[d]	95% CI
Do not report if think case has been reported in another state	27	(78)	264	(29)	2.7	2.0–3.5
Do not feel responsible for reporting	27	(63)	282	(23)	2.7	1.9–3.9
Do not think required medical information is on chart	27	(37)	282	(17)	2.2	1.3–3.8
Practice in the three largest urban counties	29	(79)	280	(48)	1.6	1.3–2.1
Do not report if think case has already been reported	27	(85)	260	(54)	1.6	1.3–1.9
Think case report forms too long	27	(26)	267	(18)	1.5	0.7–2.9
Have discrimination concerns	28	(46)	286	(39)	1.2	0.8–1.8
Have confidentiality concerns	28	(61)	286	(54)	1.1	0.8–1.6
Have liability concerns	28	(32)	284	(34)	1.0	0.6–1.7
Work in hospitals with						
>300 beds	14	(48)	149	(53)		
101–300 beds	11	(38)	84	(30)	0.7	0.4–1.6
≤100 beds	4	(14)	46	(17)	1.1	0.4–3.1

[a]Seventeen physicians did not respond to these questions; 9 indicated they report AIDS cases sometimes (25% to 75% of the time).
[b]"Rarely or never" means they report AIDS cases less than 25% of the time.
[c]"Usually or always" means they report AIDS cases 76% to 100% of the time.
[d]PR = Prevalence ratio of physicians who rarely or never report AIDS cases vs those who usually or always report.

had participated in the care of at least one person with AIDS. The numbers of AIDS patients cared for by physician specialty are listed in Table 1.

Seventy-nine percent of the physicians and 80% of the ICPs indicated that HIV infection as well as AIDS should be reportable by name to the health department. Only 28 physicians (8%) indicated that HIV infection should not be reportable by name.

Fifty-nine percent of the physicians and 63% of the ICPs indicated that primary care physicians have the principal responsibility for reporting AIDS cases. There was no significant difference in responses to this question from physician specialists and primary care physicians. Only 7% of the physicians and 17% of the ICPs indicated that ICPs have the primary responsibility for reporting AIDS cases; 9% of the physicians and 10% of the ICPs indicated that the laboratory has the primary responsibility.

Eighty-eight percent of the physicians surveyed and 83% of the ICPs indicated that they report AIDS cases most of the time or always. However, only 44% of the physicians and 64% of the ICPs report AIDS cases they think may have been reported in other states.

Physicians who replied that they rarely or never report AIDS cases were more likely than physicians who replied that they usually or always do so to indicate confusion about both the need to report cases that may have been previously reported and responsibility for reporting, a belief that information required for reporting is often not on the chart, and urban residence (Table 2). However, physicians who reported that they rarely or never report AIDS cases were not more likely to report concerns about confidentiality, legal liability associated with reporting, potential discrimination, length of the AIDS case report form, and hospital bed size (Table 2).

We used a backward stepwise logistic model to present the characteristics that are the best predictors of physicians who would rarely or never report AIDS cases (see Table 3).

Discussion

It is evident from this survey that a substantial percentage of physicians and ICPs in South Carolina have cared for AIDS patients. Similar trends may be observed in other states with low to moderate AIDS prevalence as the highest annual percent increases in AIDS cases continue to be seen in the smaller communities away from urban centers.[9] Therefore, AIDS surveillance in the less urban states is becoming increasingly important.

According to most physicians and ICPs surveyed, physicians have the principal responsibility for reporting. In South Carolina, at the time of this study, 55% of

AIDS cases were reported by physicians, 14% by ICPs, 11% by death certificate review, 6% by health department medical record review, and 14% by other means. Other authors have also indicated that practicing physicians are the key to effective surveillance of infectious diseases.[10]

Surveillance has focused on AIDS, but HIV reporting is becoming more important as preventive therapies are developed for infected patients and as health departments are becoming more involved in the clinical follow-up, case management, and funding of HIV-related disease. Health planners need to know the number and characteristics of individuals infected with HIV to plan for the cost and distribution of these services. In addition, HIV reporting has helped the development of quality partner notification and CD4 lymphocyte staging services for patients in South Carolina.[11–13]

The findings of the present study have helped DHEC surveillance personnel educate physicians and ICPs about AIDS reporting and further improve gains in active AIDS surveillance.[14] □

TABLE 3—Characteristics That Are the Best Predictors of Physicians Who Would Rarely or Never[a] Report AIDS Cases (n = 254)

Characteristics of Nonreporters	Estimate of PR[c]	95% CI
Do not report if think case has been reported in another state	16.9	2.1–134.7
Practice in the three largest urban counties	4.4	1.3–15.2
Do not think required medical information is on chart	3.8	1.4–10.7
Do not feel responsible for reporting	3.8	1.3–10.9
Have discrimination concerns	3.0	1.1–8.5

[a]"Rarely or never" means they would report AIDS cases less than 25% of the time.
[b]Respondents who did not answer all the questions included in the model (n = 94) were excluded from this analysis.
[c]Estimate of prevalence ratio (PR) of physicians who rarely or never report AIDS cases vs those who usually or always report, based on odds ratio calculated with logistic regression.

Acknowledgments

We wish to thank Ms. Julie Lumpkin for her review and excellent technical advice.

References

1. Sherman IL, Langmuir AD. Usefulness of communicable disease reports. *Public Health Rep.* 1952;67:1249–1257.
2. Langmuir AD. The surveillance of communicable diseases of national importance. *N Engl J Med.* 1963;268:182–192.
3. Chorba TL, Berkelman RL, Safford SK, Gibbs NP, Hull HF. Mandatory reporting of infectious diseases by clinicians. *JAMA.* 1989;262:3018–3026.
4. Galbraith NS. Infectious disease control. *Br Med J.* 1990;300:1477–1478.
5. Ryan White CARE Act, Pub L No. 101–381 (1990).
6. Centers for Disease Control. *HIV/AIDS Surveill Rep.* July 1989.
7. Dean AD, Dean JA, Burton AH, Dicker RC. *Epi Info, Version 5: A Word Processing, Database, and Statistics Program for Epidemiology on Microcomputers.* Stone Mountain, Ga: USD, Inc; 1990.
8. SAS Institute Inc. *SAS User's Guide: Sugi Supplemental Library Users Guide, Version 5.8.* Cary, NC: SAS Institute Inc; 1985.
9. Centers for Disease Control. Update: acquired immunodeficiency syndrome, United States, 1989. *MMWR.* 1990;39:81–86.
10. Thacker SB, Keewhan C, Brachman P. The surveillance of infectious diseases. *JAMA.* 1983;249:1181–1185.
11. Wykoff RF, Heath CW Jr, Hollis SL, et al. Contact tracing to identify human immunodeficiency virus in a rural community. *JAMA.* 1988;259:3563–3566.
12. Jones JL, Wykoff RF, Hollis SL, et al. Partner acceptance of health department notification of HIV exposure, South Carolina. *JAMA.* 1990;264:1284–1286.
13. Jones JL, Ball R, Kettinger L. The South Carolina Department of Health and Environmental Control's response to the AIDS epidemic. *J S C Med Assoc.* 1991;87:94–95.
14. Jones JL, Rion P, Hermann P, et al. Improvement in AIDS case reporting, South Carolina. *JAMA.* 1990;265:356. Letter.

Education

Physician education consisted of a grand round lecture covering the causes, evaluation, and management of falls. Community nurses were given a similar lecture emphasizing how to perform regular assessments for fall hazards in the home and recommend appropriate environmental modifications.

The thrust of our lay education was to generate awareness about falls, the importance of reporting them to appropriate health care professionals, and making home environments safer. We published articles in the community newspaper on falls and prevention. A health educator from the hospital was trained to lecture at senior centers throughout the county; a set of slides showing common hazards in the home accompanied the talks. These sessions incorporated a discussion period that allowed older people to ask for further information on falls and advice on how to make their homes safer.

Identification

To help identify fallers and those at risk, physicians and nurses were trained to question patients about falls during routine examinations. Hospital emergency room personnel were encouraged to use E codes, a subset of the ICD-9-CM codes specifically developed to identify the external cause of fall injuries, and to refer persons for physician or nurse follow-up.

During community lectures at senior centers, the health educator routinely asked whether any older persons in the audience had fallen in the past few months, took names of those people who responded affirmatively, and referred fallers to their primary providers for evaluation.

Community Services

A device bank was initiated to collect (from the community), store, and dispense unused or discarded devices (e.g., grab bars, toilet risers, bath seats) for people unable to afford them. A service was set up in which volunteers from community organizations helped older people make any necessary environmental modifications (e.g., installation of devices such as grab bars and elimination of hazardous conditions) in their homes.

Evaluation

After 1 year, we evaluated the program's effectiveness. Both physicians and nurses in the community stated that they had a greater understanding of falls. They perceived that identifying persons with falls and correcting the underlying factors responsible decreased fall risk and enhanced mobility. A review of 28 persons who had been evaluated for multiple falls verified this impression. Many suffered from several underlying medical problems (e.g., uncontrolled hypertension, electrolyte imbalance, parkinsonism, adverse drug effects), and many cited environmental hazards that were responsible for their falling episodes. Once these problems were addressed and corrected, falls were reduced. We have followed this group over a 2-year period, interviewing them on a monthly basis. To date, 75% have experienced no additional falling episodes.

The response from the broader community of older persons has been equally positive. Approximately 1500 persons have been reached at the senior centers. They generally report that the lectures have given them greater knowledge about falls and ability to identify potential fall hazards in their home environments. Although the device bank and home modification program are still in an early stage, people have appreciated the availability of devices and assistance in their installation.

This project can serve as a model of an effective low-cost fall prevention program. With the exception of one-time only grant support ($20 000) for a gerontologist to help develop the program, the costs were absorbed by the hospital (e.g., the educational slide program at a cost of $100) or were negligible. Existing resources (health professionals, community services) were used to develop prevention interventions. Currently, the program is being managed by the health educator, who has replaced the gerontologist. The health educator's salary is paid by third-party reimbursement fees generated from patient evaluations.

The acceptance and usefulness of the program are clearly demonstrated by the fact that community health professionals and organizations (e.g., Public Health Department, Visiting Nurse Service, Area Agency of Aging) have embraced it as a fall prevention measure. The hospital has incorporated the program into its ongoing budget and mainstream services offered. □

Rein Tideiksaar, PhD

Acknowledgments

This project was supported by New York State Department of Health Rural Diversification Grant C-003792. The author thanks Martin Richman, Executive Director; Marc Mendelsohn, Director of Planning; Barbara Hallenbeck, Administrator of the Geriatric Program; Peter Notarstefano, Coordinator of the Geriatric Program, Community General Hospital, Sullivan County, New York; and Judith Howe, Administrator, Ritter Department of Geriatrics and Adult Development, Mount Sinai Medical Center, New York, for their assistance in this project.

Requests for reprints should be sent to Rein Tideiksaar, PhD, Falls and Immobility Program, Ritter Department of Geriatrics and Adult Development, Mount Sinai Medical Center, One Gustave Levy Place, New York, NY 10029.

Copies of materials, including the baseline and follow-up instruments, and information about the educational materials can be obtained from the author.

Editor's Note. See related editorial on p 785 of this issue.

Community-Based Case Management of HIV Disease

Fairfax County, Virginia, with a population of more than 800 000, is the most populous political subdivision of the Washington, DC, area. Although metropolitan Washington has consistently ranked fifth nationally in the cumulative incidence of AIDS cases, Fairfax County's share of the total has been only slightly more than 5%, over 80% of whom have been classified as belonging to the gay-bisexual transmission group.

To monitor the incidence of human immunodeficiency virus (HIV) infection in its clinic populations, the Fairfax County Health Department began an expanded program of voluntary testing for antibody during 1986. In June 1987 the county implemented the HIV Case Management Program for residents asymptomatically infected with HIV. The program was designed primarily to serve the low-income, high-risk heterosexual populations, especially those who have engaged in intravenous drug use, and their sexual partners. These individuals are difficult to reach and may pose the greatest danger for continued spread of the virus in the general population. The program is open to the gay-bisexual community as well, but because most have access to other care providers through insurance or special community resources, many opt for care through other sources.

The county's human services agencies—Department of Health, Fairfax-Falls Church Community Services Board, and the Department of Human Development—operate this program jointly. During the past year, by contractual arrangement, a group practice specializing in infectious diseases (Infectious Diseases Physicians, Inc. or IDP) has provided a program for early

TABLE 1—Case Management Program Costs by Agency for Calendar Years 1989 and 1990

	1989 (33 visits, 23 patients)	1990 (45 visits, 36 patients)
Health Department		
Staff	$25 371	$27 823
Laboratory	$ 5555	$ 9000
Subtotal	$30 926	$36 823
Community Services Board	$ 6376	$10 464
Human Development	$ 3125	$ 2235
Total	$42 427	$49 522

Zidovudine therapy for asymptomatic HIV patients. The county health department collaborates with IDP in monitoring these patients until they become symptomatic or require prophylaxis against *Pneumoncystis carinii* infection. At that stage in the disease, patients are no longer followed in the case management program. Fairfax Hospital, under a contract with the county health department, assumes medical and case management responsibility for our patients when their disease becomes symptomatic or their CD4 count falls below 200.

The case management program holds clinics at two geographic sites within the county; the clinics are staffed by personnel from the three participating agencies. The management of patients with HIV infection includes lifestyle counseling to slow the progression of the infection and to prevent its spread to others. A special patient consent form enables members of the interagency team to share patient information with each other, thus facilitating the development of a coordinated plan for the patient's medical care and mental health and social needs. This systematic plan reduces the need for crisis management.

Patients enter the case management program voluntarily. They are referred to the program through the health department's HIV screening and case-finding activities and by private physicians, local hospitals, and HIV/AIDS information services. Before the first clinic visit, the patient provides a detailed medical and social history and blood and urine specimens for routine laboratory examination.

At the initial and subsequent routine visits the patient's history is updated and a thorough clinical examination is performed. Routine laboratory tests include complete blood cell count and differential, urinalysis, serologic tests for syphilis, HIV antigen test, β-2 microglobulin assay, and a helper T lymphocyte (CD4) count. Tests for hepatitis B core antibody and surface antigen are done on the initial visit and subsequently if indicated. Pap smears are done annually on female patients. Chest x-rays are obtained for the initial visit and annually thereafter. Reactivity to purified protein derivative (tuberculin), mumps, and candida antigens is determined by skin test at the initial visit.

Patients are seen every 6 months unless a shorter interval is indicated; those receiving early Zidovudine therapy are seen at intervals set by IDP. At each routine visit the patient is evaluated and counseled by the mental health counselor and, if indicated, the substance abuse counselor. Arrangements are made for separate individual or group counseling sessions as desired. The Human Development team member reviews with the patients their current and future financial requirements and needs for support. Case management program team members meet jointly once a month to review the status of the patients seen during the month and to make recommendations for their future care.

A total of 81 patients entered the case management program between June 1987 and June 1991. Demographic attributes of the patients are as follows:

- Age: Range, 21 to 62 years; median, 30 years
- Sex: Male 51, female 30
- Ethnic group: White, 38; Black, 39; Hispanic, 3; Asian, 1
- Risk group: Male homosexual, 27; intravenous drugs, 34; heterosexual, 20

The characteristics of the 14 patients lost to follow-up were similar to those of the total group. Only one patient, an elderly person with chronic alcoholism, is known to have died. The cause of death was liver failure and HIV was not considered a contributory cause.

Program costs rose only slightly from 1989 to 1990, whereas costs per visit declined (Table 1). The total cost per visit in 1989 was $1285; in 1990, it was $1100. The average for the 2 years was $1179 per visit.

The health department's cost for routine clinic services in 1990 was approximately $400 per visit, which included $200 for laboratory services. The remainder of the staff costs were devoted to intake, patient follow-up, and case review. The increase in laboratory costs in 1990 reflects the greater use of CD4 counts for patients receiving early Zidovudine therapy. Most clinics operated at considerably less than optimum patient load; the operating costs per visit should decline appreciably with increasing numbers of patients per clinic.

The collaborative program with IDP for early therapy with Zidovudine entails additional annual costs of $400 per patient for physicians' fees. At present the state of Virginia is supplying the Zidovudine if the patient is unable to pay for the drug; thus this cost is not included in the direct costs of the program.

Although the future course of the HIV epidemic in the United States is still uncertain, current evidence suggests that the virus has begun to spread in the young heterosexual population in many areas of the country. There will be an increased need for adaptive programs at the community level that will combine expanded case-finding activities with case-management programs similar to the one described here.

As the next step in this program, we are exploring new methods of outreach to intravenous drug users and to the young sexually active heterosexual population. Improved methods will be required if there is to be effective intervention in this epidemic. Programs like this can provide a useful community focus for these broader efforts. □

Fred J. Payne, MD
Carol S. Sharrett, MD
Donald N. Poretz, MD
Lawrence J. Eron, MD
Thomas Stage, MD
Rosalyn Foroobar, RN, MA
Chester Bowman, BS
Richard K. Miller, MD

Fred J. Payne, Carol S. Sharrett, Rosalyn Foroobar, Chester Bowman, and Richard K. Miller are with the Fairfax County Health Department, Fairfax, Va. Donald N. Poretz and Lawrence J. Eron are with the Infectious Disease Section, Department of Medicine, Fairfax Hospital, Fairfax, Va. Thomas Stage is with the Community Services Board, Fairfax, Va.

Requests for reprints should be sent to Fred J. Payne, MD, Director, HIV Case Management Program, 10777 Main Street, Suite 203, Fairfax, VA 22030.

A2623-1

American Journal of Public Health

ED KIANKA
CONGRESSIONAL INFORMATION SVC
4520 EAST-WEST HIGHWAY
BETHESDA MD 20814
86389
028

July 1992, Vol. 82, No. 7 Established 1911

Editorial

931 Determinants of Worldwide Health *W. Winkelstein*

Public Health Policy Forum

933 Health and the Environment in the 1990s *R. Doll*

941 Commentary: Environmental Disease—A Preventable Epidemic *P. J. Landrigan*

Featuring Environment/Occupational Health

944 The Potential Impact of Epidemiology on the Prevention of Occupational Disease *D. H. Wegman*

955 Chlorination, Chlorination By-products, and Cancer: A Meta-analysis *R. D. Morris, A.-M. Audet, I. F. Angelillo, T. C. Chalmers, and F. Mosteller*

964 Epidemiological Data on US Coal Miners' Pneumoconiosis, 1960 to 1988 *M. D. Attfield and R. M. Castellan*

971 Surveillance Data on US Coal Miners' Pneumoconiosis, 1970 to 1986 *M. D. Attfield and R. B. Althouse*

978 British Data on Coal Miners' Pneumoconiosis and Relevance to US Conditions *M. D. Attfield*

984 Job Strain and Prevalence of Hypertension in a Biracial Population of Urban Bus Drivers *C. L. Albright, M. A. Winkleby, D. R. Ragland, J. Fisher, and S. L. Syme*

990 Use of Hair Coloring Products and the Risk of Lymphoma, Multiple Myeloma, and Chronic Lymphocytic Leukemia *S. H. Zahm, D. D. Weisenburger, P. A. Babbitt, R. C. Saal, J. B. Vaught, and A. Blair*

1011 Rotating Shift Work, Sleep, and Accidents Related to Sleepiness in Hospital Nurses *D. R. Gold, S. Rogacz, N. Bock, T. D. Tosteson, T. M. Baum, F. E. Speizer, and C. A. Czeisler*

1014 Birthweight Distributions in Mexico City and among US Southwest Mexican Americans: The Effect of Altitude *F. C. Notzon, J. L. Bobadilla, and I. Coria*

(continued page 921)

APHA
American Public Health Association

Featuring Environment/ Occupational Health

Contents

American Journal of Public Health

The Journal's one-page guidelines for authors, "What *AJPH* Authors Should Know," are printed in each issue. They can also be obtained from the Editorial Office.

American Public Health Association

July 1992, Vol. 82, No. 7 ISSN: 0090-0036 Established 1911

Articles

999 Education, Race, and High-Density Lipoprotein Cholesterol among US Adults *D. S. Freedman, D. S. Strogatz, D. F. Williamson, and R. E. Aubert*

1007 Young Women's Antisocial Behavior and the Later Emotional and Behavioral Health of Their Children *S. L. Martin and M. R. Burchinal*

Public Health Briefs

1018 Identification of Nonfatal Family and Intimate Assault Incidents in Police Data *L. E. Saltzman, J. A. Mercy, and P. H. Rhodes*

1020 The Cost and Frequency of Hospitalization for Fall-Related Injuries in Older Adults *B. H. Alexander, F. P. Rivara, and M. E. Wolf*

1023 Carcinogens in Tobacco Smoke: Benzo[a]pyrene from Canadian Cigarettes and Cigarette Tobacco *M. J. Kaiserman and W. S. Rickert*

1026 Discrepancies between Self-reported Smoking and Carboxyhemoglobin: An Analysis of the Second National Health and Nutrition Survey *L. M. Klesges, R. C. Klesges, and J. A. Cigrang*

1029 Recall of AIDS Public Service Announcements and Their Impact on the Ranking of AIDS as a National Problem *M. Siska, J. Jason, P. Murdoch, W. S. Yang, and R. J. Donovan*

Public Health Then and Now

1033 The President's Committee on Health Education: A 20-Year Retrospective on Its Politics and Policy Impact *M. A. Guinta and J. P. Allegrante*

Other Journal Departments

922 APHA Masthead

928 Advertisers' Index

930 What *AJPH* Authors Should Know

931 Journal Masthead

998 Erratum

1042 Letters to the Editor

1045 Book Corner

1050 Annual Meeting Registration Form and Information

1055 Job Opportunities

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.

APHA OFFICERS
Joyce C. Lashof, MD, *President*
C. William Keck, MD, MPH, *Immediate Past President*
Helen Rodriguez-Trias, MD, *President-Elect*
Arnold E. Greenberg, SM, *Vice-President (USA)*
Trevor Hancock, MB, MHSc, *Vice-President (Canada)*
Julio Teja Perez, MD, *Vice-President (Latin America)*
Caswell A. Evans, Jr, DDS, MPH, *Treasurer*
Frances M. Veverka, RS, MPH, *Chair, Executive Board*
George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*
William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD
Frances M. Veverka, RS, MPH (1992) *Chair*
Marilyn Aguirre-Molina, EdD, MS (1993)
Brian A. Dobrow, MPH (1992)
Caswell A. Evans, Jr, DDS, MPH, *Treasurer*
Kristine M. Gebbie, BSN, MN (1994) *Vice-Chair*
Alan R. Hinman, MD, MPH (1995)
Janet Heinrich, DrPH, MPH (1993)
Margaret H. Jordan, RN, MPH (1992)
C. William Keck, MD, MPH, *Immediate Past President*
Margo D. Kerrigan, MPH (1995)
Joyce C. Lashof, MD, *President*
Barry S. Levy, MD, MPH (1993)
Helen Rodriguez-Trias, MD, *President-Elect*
Allan G. Rosenfield, MD (1994)
Reed V. Tuckson, MD (1994)
Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO
William H. McBeath, MD, MPH, *Executive Director*
Joanne E. Lukomnik, MD, *Chair, Action Board*
George Pickett, MD, MPH, *Chair, Program Development Board*
George E. Hardy, MD, MPH, *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc, 1015 Fifteenth St, NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50 000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health*, 1015 Fifteenth St, NW, Washington, DC 20005.

Information on Manuscript Preparation is published monthly in the Journal. The Journal publishes reports of original research, research methods, program evaluations, analytic reviews, including health policy analysis, and reports for special departments. A paper is considered with the understanding that it or its essential substance neither has been published nor is under consideration elsewhere.

Subscription price, individual, is $80/year US, $120/year international; **institutional,** $160/year US, $240/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to *American Journal of Public Health*, 1015 Fifteenth St, NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St, NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St, NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS/COLLEAGUE, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc, Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Public Health Policy Forum

Health and the Environment in the 1990s

Richard Doll, DM, FRS

ABSTRACT

Expectation of life at birth provides a simple measure of the state of health of a country. Differences in the expectation are examined in the United States over time and between males and females and Whites and non-Whites, and worldwide between market and nonmarket developed countries, and between developed and developing countries. The principal factors responsible for the trends and the current differences are changes in the social and physical environment, in personal behavior, and in medical care, and their relative importance is assessed. It is concluded that, at present, the principal environmental hazards worldwide are those associated with poverty of individuals within the market economies and of communities in the developing countries and that in the future, they will be the effects of overpopulation and the production of greenhouse gases. (*Am J Public Health.* 1992;82:933–941)

Health, according to the World Health Organization (WHO), is defined as a state of complete physical, mental, and social well-being. This is a fine and inspiring concept and its pursuit guarantees health professionals unlimited opportunities for work in the future, but it is not of much practical use for specialists in public health medicine who need to compare the states of health in different communities and at different times and who consequently need to give them numerical values. I shall therefore use the term *health* in the limited sense of a state distinguished by the absence of disease or of physical or mental defect, that is, the absence of conditions that detract from functional capacity whose incidence can be measured objectively. And because morbidity and mortality are so closely related, I shall assess health largely in terms of mortality and years of expectation of life, for which objective evidence is available for long periods throughout most of the world.

As to environment, I shall start with the simple definition of the *Oxford English Dictionary:* namely, "the conditions under which any person lives." We can, however, limit the conditions to those whose effects people are individually unable to control, and so exclude such conditions as the availability of tobacco and, in high-income societies, of a high-fat diet, as people exposed to them can choose alternative diets and whether or not to smoke. By excluding these conditions I do not exclude the effects of malnutrition in low-income societies in which a healthy diet may be beyond people's means; in these circumstances, a nutritionally inadequate diet must be regarded as a condition of the society in which the people live.

Even with these restricted definitions, the subject is still so vast that I can deal with much of it only in outline. I shall therefore describe first and very briefly the trends in mortality in the United States, as representative of what the World Bank calls the market economy countries of the industrialized world,[1] and the comparative conditions in the non–market economy industrialized countries and the developing countries. I shall examine in each case the environmental factors that have contributed to the trends and to the differences between them. Second, I shall seek to justify my statement that the changes that occur in each part of the world affect conditions everywhere else. I shall conclude by suggesting some priorities for the selection of the environmental factors that we should now most energetically seek to control.

Health and the Environment in the United States

Expectation of Life

The simplest measure of the state of health of a country is the expectation of life at birth of its citizens. The trend in expectation of life at birth in the United States since the beginning of the century is shown in Figure 1, separately for males and females. The data for the years before 1928 refer only to the limited number of registration states (10 plus the District of Columbia until 1911 and then 30 plus the District of Columbia until 1927), but the lack of sharp breaks in the trend suggests

The author is with the Imperial Cancer Research Fund Cancer Studies Unit, Radcliffe Infirmary, Oxford, England.

Requests for reprints should be sent to Sir Richard Doll, DM, FRS, Imperial Cancer Research Fund Cancer Studies Unit, Radcliffe Infirmary, Oxford OX2 6HE, UK.

This paper was submitted to the Journal February 20, 1992, and accepted March 24, 1992.

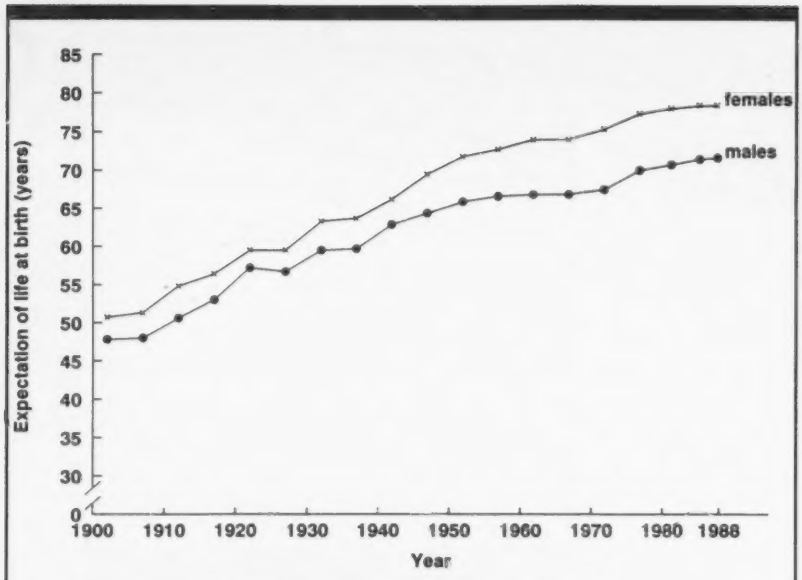


FIGURE 1—Trend in expectation of life at birth in the United States, 1900 to 1985, for males (●——●) and females (x——x).

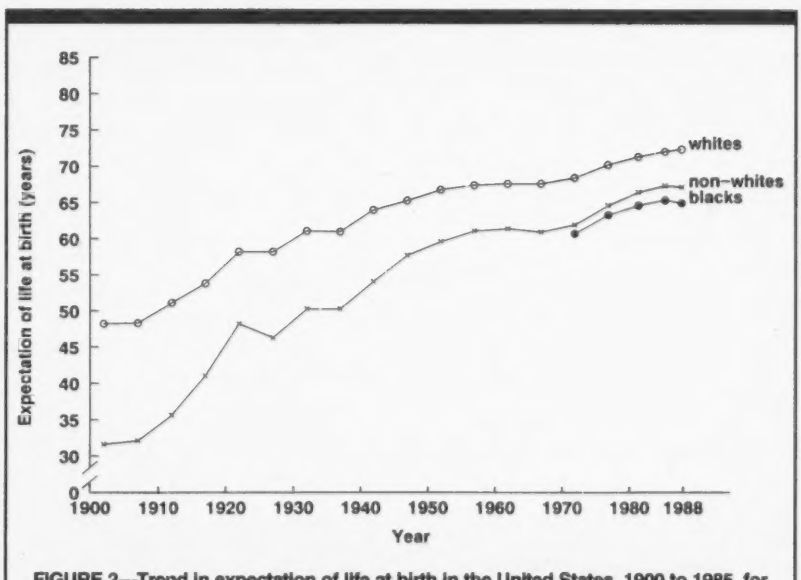


FIGURE 2—Trend in expectation of life at birth in the United States, 1900 to 1985, for Whites (○——○), non-Whites (x——x), and Blacks (●——●).

that the pre-1928 figures were reasonably representative of the whole country. In both sexes increases have occurred throughout the period, apart from brief intervals in the late 1920s, 1930s, and 1960s. In total the increases have been large. The increase has been greater for females than for males, both absolutely and proportionately, and the female advantage of nearly 3 years in 1900 to 1904 has more than doubled to nearly 7 years in the late 1980s.

In all industrialized countries the state of health varies between social groups, and this has been particularly easy to see in the United States because of the publication of vital statistics by race and the broad correlation between race and social status. The differences in the expectation of life between White and non-White males at different times in this century are shown in Figure 2. The 1900 advantage of nearly 17 years for White males over non-White males (almost all of whom at that time were Black) has progressively diminished, but in 1988 the difference was still about 5 years or, if comparison is made only with Blacks, for whom separate data are now published, about 7 years.

Factors Responsible for Trends

These increases in the expectation of life have been brought about primarily by a reduction of over 90% in the mortality rate for infants and children younger than 5 years of age, which was 30% at the start of the century and is now only 2%. As a result, the increase of 23 years in the expectation of life at birth corresponds to an increase of only 14 years for those who survive to reach the age of 5. Reductions in mortality have, however, occurred at all other ages, as is shown in Figure 3; even for those older than 80 years of age the reduction has been as much as a third.

The reasons for the reduced mortality in the early decades of the century are well known and not open to serious question, except as regards the relative importance of the various factors responsible for the reduction. These were essentially factors that diminished the risk or the fatality of infection, including improved nutrition, smaller family size, better provision of water uncontaminated with feces, control of the vectors of malaria, pasteurization of milk, maternal education in the care of infants, and immunization. Treatment of infection played little if any part in the reduction of mortality, as treatment was generally ineffective until the introduction of the sulfa drugs in 1936. The principal roles in reduction of mortality must be attributed to the improved environment (defined as the conditions under which people lived) brought about by the increased standard of living consequent on industrial development and, secondarily, to the application of knowledge of the causes of disease and the way disease could be prevented. With the advent of sulfa drugs and the development of a series of antibiotics following the first use of a laboratory preparation of penicillin in 1941, therapy has played an increasingly important role in the control of infection and, more recently, in the treatment of cancer and ischemic heart disease. Other factors are better anesthesia and postoperative care, further immunization, and the discovery of the means to reduce substantially the age-specific risk of degenerative disease by the avoidance of smoking and the reduction of blood pressure and of blood low-density-lipoprotein cholesterol.

The longer expectation of life of females is a universal phenomenon in industrialized countries; it is partly genetic in origin and partly behavioral. The widening of the gap between the sexes that began immediately after the Second World War and continued into the 1970s, and its subsequent narrowing (Table 1), can be attributed largely to cigarette smoking being taken up first by men and only later by women. The different trends in mortality from the principal diseases related to smoking (heart disease, respiratory cancer, and chronic obstructive pulmonary disease) could account for most of the difference between the trends in the two sexes from 1950 to 1970 and for all of it from 1970 to 1988 (Table 2).

The factors responsible for the shorter expectation of life of Black Americans than of Whites and for the partial closing of the gap up to 1960 are equally clear, though again difficult to put in hierarchical order. In this case, the direct medical effects of genetic factors operate in both directions, and essentially the whole difference must be attributed to differences in the social environment in which people live, operating sometimes indirectly by modifying behavior and sometimes directly via malnutrition, overcrowding, and a deficiency of medical care.[2-4] The narrowing of the gap, which occurred between 1900 and 1960, can be attributed almost entirely to the differential reduction in mortality in infancy and childhood, as the difference in the expectation of life between Whites and Blacks in adult life has changed relatively little (Table 3). The narrowing of the gap, therefore, was due to the factors referred to previously that improved the health of infants and children in all sections of the community, principally by reducing the risk of infection.

Role of the Environment

These factors have their origins partly in increased scientific knowledge, but at least as much, if not much more, in the changes in the environment that have accompanied the increased wealth produced by industrial and agricultural development, that is, the reduced prevalence of malnutrition and of imperfectly preserved food, the provision of ample water supplies uncontaminated with pathogenic organisms, the reduced prevalence of pathogenic parasites or their vectors, the better housing and smaller families, and the increased opportunities for education. There are, however, other aspects of industrial and agricultural development that have had deleterious effects, for example, the destruction of much natural beauty; the loss of open spaces for exercise; the destruction of natural habitats for many species, leading, in some cases, to their extinction; the widespread distribution of potentially toxic chemicals; and the distribution of other chemicals in nontoxic amounts but on such a scale that the climate of the world is modified. I shall consider the last effect later, as it is its potential for future damage that is most worrying, but we need to consider now how far the dispersal of toxic chemicals may already have harmed health and partially counterbalanced the many benefits of increasing prosperity.

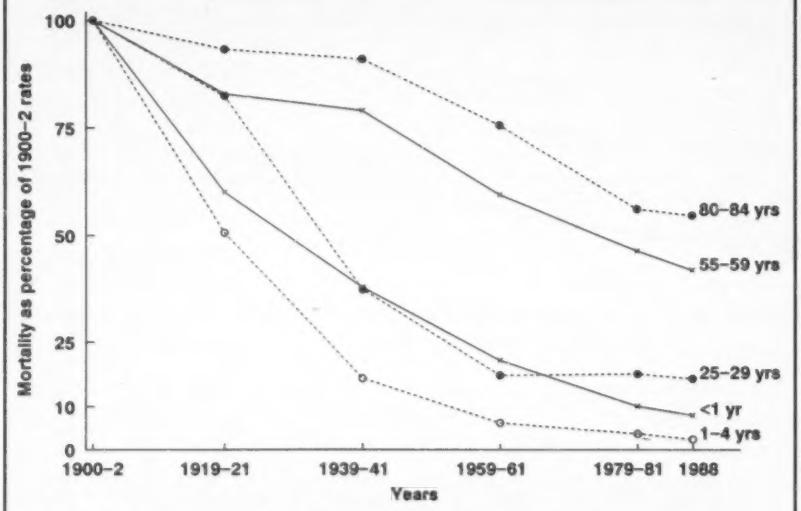


FIGURE 3—Trend in mortality by age in the United States, 1900–1902 to 1988, ages less than 1 year (x——x), 1–4 years (○—○), 25–29 years (●—●), 55–59 years (x——x), and 80–84 years (●—●).

TABLE 1—Expectation of Life at Birth, by Sex: 1900–1924 to 1985–1988

Period	Expectation of Life, y		
	Male	Female	Female Advantage
1900–1924	51.9	55.0	3.1
1930–1944	61.4	64.8	3.4
1950–1954	65.9	71.8	5.9
1960–1964	66.8	73.5	6.7
1970–1974	67.5	75.2	7.7
1980–1984	70.7	77.9	7.2
1985–1988	71.4	78.3	6.9

This problem has given rise to a vast literature, but hard evidence of the material effects on health of the dispersal of toxic chemicals is relatively sparse. What is clear is that industrial developments over the last century have affected health in two ways. First, the pollution in urban air has been reduced—enormously in the case of smoke and substantially in the case of sulfur dioxide—by the use of central heating instead of domestic fires; by the use of natural gas, oil, and electricity as sources of power, instead of coal and wood; and by the control of emissions when coal is burned as a source of energy in industry. Second, complex pollution has been caused by the enormous increase in motor traffic and its accompanying exhaust fumes; by the wide dispersal of natural and artificial chemicals in emissions from factories; by the use of chemical fertilizers, herbicides, and pesticides in agriculture; and by the dispersal of industrial waste. What is less clear is whether, and to what extent, the harm done by the increase in complex pollution outweighs (if it does) the benefits of the reduction in urban air pollution.

The difficulty is that few of these effects have been properly quantified. The pollution of the past (which is still present in the non-market economy countries and in some of the developing countries) caused dense smogs when accompanied by temperature inversions, and these smogs occasionally caused acute increases in mortality from heart and respiratory diseases. One such smog in London in 1952 was estimated to have caused 4000 premature deaths,[5,6] but this was exceptional and the normal annual total pro-

TABLE 2—Mortality from Different Causes, by Sex: 1950 to 1988

Cause of Death	Sex	Mortality per 100 000	Percent change	
		1950	1950–1970	1970–1988
Heart disease, respiratory cancer, and chronic obstructive pulmonary disease	M	498.7	+3.1	−27.3
	F	291.0	−20.7	−16.2
Other	M	554.4	−14.9	−24.1
	F	414.0	−23.1	−27.5

TABLE 3—Expectation of Life at Different Ages, by Race: 1900 to 1981

Age	Period	Expectation of Life, y		
		White	Black	White Advantage
Newborn	1900–1902	49.7	34.1	15.6
	1959–1961	70.9	64.0	6.9
20 years	1900–1902	43.0	36.0	7.0
	1950–1961	53.3	47.9	5.4

duced in this way may have been quite small. The very high concentrations of smoke—which in large towns commonly averaged over 100 $\mu g/m^3$ of particulate matter throughout the year, with a content of 30 to 40 ng/m^3 of benzo(a)pyrene—must have produced some lung cancers, but direct evidence of the size of the effect has never been clearly obtained, particularly for nonsmokers. Extrapolation from industrial exposures suggests that urban exposure may have contributed, in conjunction with other polycyclic hydrocarbons in the air, to some 10% of the lung cancers in populations in cities with average smoking habits.[7,8] If this were so, the reduction in smoke pollution to levels one or two orders of magnitude less could have prevented many hundreds of premature deaths a year.

How much ill health has been produced by the new sources of pollution is even more difficult to estimate. That some has been produced is incontrovertible: for example, the irritation of the eyes and respiratory system by the photochemical smog produced by the action of sunlight on the exhaust fumes of cars, the damage to the developing brain from environmental lead, and a few cases of mesothelioma from the dispersal of asbestos. But whether current levels of sulfur dioxide in combination with nitrogen oxides and the corresponding acids are responsible for more respiratory disease than were the previous higher levels of sulfur dioxide in combination with smoke is impossible to assess directly, as is the total effect on public health of the dispersal of industrial products and waste into air, soil, and water.

One particular type of pollution that has caused much public concern is the pollution of food with residual carcinogenic pesticides. Yet Ames and Gold[9] estimate that the amount of such carcinogens in food is 1800 times *less* than the amount of carcinogens derived from the 54 natural plant chemicals in food that have been similarly tested for carcinogenicity. Another cause of public concern is the dispersal of hazardous industrial waste. This subject has been considered by a Committee of the Boards on Environmental Studies and Toxicity of the National Research Council. The committee concluded that many criticisms could be made of the 21 published reports of the effects on public health of hazardous waste sites, but that the evidence was strong enough in sum to suggest that "hazardous waste sites have produced serious health effects in some populations." On the basis of its review of the published literature, however, the committee was unable to answer the question "whether Superfund and other hazardous-waste programs actually protect human health." Yet "less than 1% of the estimated $4.2 billion spent each year on hazardous waste sites in the United States [during the last 10 years] has been used to evaluate health risks at listed Superfund sites."[10]

Irrespective of the conclusion that may be reached about individual sources of pollution, it is clear that modern industrial and agricultural development has not caused a new epidemic of either of the most serious effects that have most commonly been postulated, namely cancer and congenital anomalies.

The increase in the mortality from cancer at all ages that is sometimes cited as evidence of increasing environmental hazards is due to an increase at old ages and can be accounted for in all industrialized countries by the spread of cigarette smoking by men in the middle of the century and by women more recently. When this effect is allowed for properly, the mortality from all other cancers, considered as a group, is decreasing in middle and old age.[11] To detect an effect from new hazards of individual types of cancer we need to examine the trends in incidence and mortality in early middle age (that is, in adults under 50 years of age), as it is in the relatively young that the effect of new hazards is usually first seen. Examination of these trends shows that most of the important cancers are becoming less common in both Britain and the United States, the increase in the recorded incidence of the others is mostly an artifact of screening or the result of behavior that the individual can choose to avoid.[12,13] Only the worldwide increase in cancer of the testis and the small increase in non-Hodgkin's lymphoma that is not attributable to immune deficiency remain unexplained and may possibly be due to new environmental hazards.

Less evidence is available for congenital malformations, and that which is available is difficult to interpret. When changes in prevalence do occur, the evidence—according to Knox and Lancashire,[14] who have made a detailed examination of the trends observed in a large British series between 1964 and 1984—points firmly to changes in case-finding or to infection rather than to toxic agents. In the United States two registers of malformations are organized by the Centers for Disease Control in Atlanta. One covers about 35 000 births a year in metropolitan Atlanta and seeks information from a variety of sources for the first year of the child's life. The other covers about 575 000 births a year in 1200 hospitals throughout the country and seeks information from computerized discharge diagnoses. Both have operated since 1974. Table 4 shows the results obtained for five types of abnormality, which should be easily diagnosed at birth (cleft lip, cleft palate, reduction deformity of limb, esophageal atresia, and anal atresia).[15] The drop in prevalence between 1979 and

1980 in the Atlanta registry was principally due to a decrease in the prevalence of cleft lips and palates, which does not appear to be an artifact of registration and has yet to be explained (M. J. Khoury, personal communication). The larger register, in contrast, showed a small increase in several of these malformations in the early 1980s, which might have been due to a small real increase in prevalence, to improved registration, or to a change in the nature of the hospital populations.

TABLE 4—Trend in Prevalence of Selected Congenital Malformations, 1974 to 1988

Register	Prevalence per 10 000 Infants[a]				
	1974–1976	1977–1979	1980–1982	1983–1985	1986–1988
Atlanta	31.9	30.2	25.9	26.7	23.0
US hospitals	22.5	21.6	22.7	24.8	24.4

[a]Stillbirths of 20 weeks or more gestation are included.

TABLE 5—Expectation of Life in Industrialized Countries, circa 1988

Type of Economy	Expectation of Life at Birth, y		
	Minimum	Median	Maximum
Market (23 countries)	70.7 (Finland)	73.1	76.2 (Japan)
Non-market (8 countries)[a]	64.7 (USSR)	67.4	70.1 (GDR)

Note. Adapted from World Health Organization.[16]
[a]Data for Albania not available.

TABLE 6—Mortality in Industrialized Countries by Type of Economy and Cause: 1985

Cause of Death	Annual Death Rate per 100 000 Persons[a]		
	Market Economy (A)	Non-market Economy (B)	B as percentage of A
Tuberculosis	1.8	7.7	427
Other infection	36.7	41.4	113
Neoplasms	209.5	201.2	96
Ischemic heart disease	165.6	311.6	188
Cerebrovascular disease	92.9	195.4	210
Diabetes	14.1	8.6	61
Complications of pregnancy	0.1	0.8	800
Perinatal conditions	6.7	12.0	179
Injury or poisoning	55.7	89.5	161
Other defined	129.0	148.1	115
Ill defined	22.6	20.2	89
All causes	853.1	1205.4	141

Note: Adapted from Lopez.[1]
[a]Death rates are averages of rates for males and females.

Health and the Environment in Other Countries

Non–Market Economy Industrialized Countries

When we look at other countries the picture is different—not so different in the non–market economy industrialized countries, but sufficiently different for the total mortality rate adjusted for age to be 41% higher in these countries than in the market economy countries and for nearly all of the cause-specific mortality rates to be higher as well. Mortality in infancy and childhood, however, is low in both types of industrialized countries, and the increased rates in the non–market economy countries affect principally the old. The expectation of life in 1988 and 1989 nevertheless tended to be about 6 years lower in the non–market economy countries than in the market economy countries,[16] as is shown in Table 5.

As Table 6 shows, the greatest relative differences in mortality between the two types of economy are for tuberculosis and complications of pregnancy, but neither contributes importantly to the total. Nearly two thirds of the absolute excess (63%) is due to ischemic heart disease and cerebrovascular disease, the mortality of which has been decreasing in Western Europe and North America as a result of prophylactic treatment and altered behavior. Some of the difference must be environmental—overcrowding, for example, contributes to the spread of tuberculosis—but how much is environmental remains a subject for research.

Developing Countries

The situation in the developing countries is very different. For many countries, precise information is lacking about both numbers of deaths and the causes of death, but estimates can be made from surveys of incidence and studies of fatality. For some others, data exist that are as good as those for the industrialized countries, with which some of the developing countries now deserve to be classified. In China, whose population constitutes nearly 30% of the total population of developing countries, the infant mortality rate per 1000 births fell from 195 in 1950 to 39 in 1985, and the expectation of life at birth increased from 40.8 years to 67.8 years—close to the median for that of the non–market economy industrialized countries. In the other developing countries, taken as a whole, infant mortality fell only from 173 to 100 per 1000 births and the life expectancy increased correspondingly less, from 41.7 years to 55.7 years.

WHO has estimated the number of deaths that occurred in developing countries in one year in or around 1985 and classified them into a few broad categories of cause for three geographical groupings.[1] In Table 7, the results are compared with similar data for the industrialized countries as a whole (market and non-market economies combined). Nearly 40 million of the 50 million or so deaths that were estimated to have occurred in a year in the whole world were in the developing countries. The biggest absolute and relative differences (other than the relative difference in deaths due to the complications of pregnancy) were in the proportion attributed to infectious and parasitic diseases and in the proportions attributed to neoplasms and to circulatory and certain

TABLE 7—Mortality by Cause as a Percentage of All Causes in Developing and Industrialized Countries, circa 1985

Cause of Death	Developing Countries: Africa and Middle East (11.5)	Developing Countries: Asia (23.2)	Developing Countries: Latin America and Caribbean (3.2)	Industrialized Countries (11.0)
Infectious and parasitic disease	60.0	39.7	28.1	4.6
Neoplasms	3.9	7.5	9.4	20.8
Circulatory and certain degenerative diseases	10.4	19.0	28.1	53.7
Chronic obstructive lung disease	1.0	9.1	2.8	3.5
Complications of pregnancy	1.8	1.1	1.1	<0.1
Perinatal conditions	9.6	7.8	9.4	0.9
Injury and poisoning	4.8	6.9	7.8	7.0
Other and unknown	8.6	9.0	13.3	9.5

Note. Adapted from Lopez.[1] Numbers of deaths in millions per year are shown in parentheses.

TABLE 8—Deaths Attributed to Infectious and Parasitic Diseases in Developing Countries, circa 1985, by Age and Cause

Cause of Death	Annual Deaths (in thousands): <5 years of age	Annual Deaths (in thousands): ≥5 years of age
Diarrheal diseases	4000	1000
Tuberculosis	300	2700
Measles,[a] whooping cough, and diphtheria	2200	...
Malaria[b]	750	250
Schistosomiasis	...	200
Acute respiratory diseases	2800	2000
Other infectious and parasitic diseases	450	350
Total	10 500	6500

Note. Adapted from Lopez.[1]
[a]400 000 measles-related deaths are included under diarrheal diseases.
[b]Up to 1 million malaria-related deaths are included under other infectious and some noninfectious diseases.

degenerative diseases. The proportion of deaths attributed to infectious and parasitic diseases *dropped* progressively from a high of 60.0% in developing countries in Africa and the Middle East through somewhat lower rates in Asia and in Latin America and the Caribbean to a low of 4.7% in the industrialized countries. The proportions of deaths attributed to neoplasms and to circulatory and some degenerative diseases *rose,* in contrast, in the same order (from 3.9% to 20.8% and from 10.4% to 53.7%, respectively).

The differences between national groups in mortality from diseases that increase in incidence with age are largely due to the different age distributions of the populations—the smallest proportion of old people is in the developing countries of Africa and the Middle East. In part, however, they also reflect differences in the patterns of disease at each age, with high incidence rates of vascular disease and neoplasms in the industrialized countries due principally to the eating, drinking, smoking, and reproductive behavior of the inhabitants.[8]

The high proportion of deaths attributed to infectious and parasitic diseases (including influenza, acute bronchitis, and pneumonia) in the developing countries falls most heavily on children younger than 5 years of age (Table 8). Over 10 million deaths a year from infectious and parasitic diseases (61% of the total) are estimated to occur in this age group; several million of these deaths are due to diseases that have been almost eliminated from the industrialized countries and entirely eliminated from some of them. The attribution of each of these deaths to a single cause (which was required for the preparation of the table) is, however, an even more arbitrary procedure than it often is for the deaths of old people in industrialized countries. Children often have multiple infections that are made severe by malnutrition and there may be little reason to choose one as the cause of death rather than another. Many of the deaths attributed to diarrheal diseases are related to measles, and up to a million deaths classified under "other infectious and parasitic diseases" and some noninfectious diseases are related to malaria. The total number of deaths due to malaria is threatening to rise above 2 million a year as chloroquine-resistant parasites spread more widely.

The deaths attributed to "other infectious and parasitic diseases" include about 25 000 attributable to poliomyelitis (80% in children younger than 5 years of age) and many thousands from amebiasis, hookworm, hepatitis B (excluding the complication of liver cancer), and acquired immunodeficiency syndrome (AIDS). A further 800 000 deaths attributed to neonatal tetanus, practically all of which occur in the developing countries, are not included in the figures for infectious disease, as they are classified under perinatal deaths.

This high mortality in developing countries, serious though it is, is lower than it was at the beginning of the century and there is no lack of knowledge to prevent it being reduced almost to the level in the industrialized world. The causes of the current excess of ill health in the developing world are for the most part environmental in origin, in the broad sense of the term, and are essentially avoidable. Head and shoulders above the rest is poverty—poverty of the individual that results in malnutrition and lack of shelter, and poverty of the state that makes it difficult to provide adequate standards of education and immunization and to provide the classical public health measures that were introduced into the industrialized countries before, or shortly after, the beginning of the century. These measures are difficult, but not impossible, to provide, as was demonstrated in China in the 20 years that followed the start of the Communist regime, when a dictatorial and bureaucratic system of social organization was able to ensure that central directives were effectively carried out, agricultural production was improved, food was more evenly distributed, an immunization program was fulfilled, and fatal childhood infections were greatly reduced.

Other environmental hazards that affect primarily adults include many that were characteristic of the industrialized world during its period of development. Occupational hazards, like that from asbestos, and general atmospheric pollution with smoke from the industrial use of coal

are often poorly controlled. More important, because it is more widespread and affects a larger proportion of the population, is the indoor pollution from the use of biomass fuels for cooking and heating in small rooms with inadequate ventilation that is prevalent in north China and parts of India. This indoor pollution may be partly responsible for the high proportion of deaths (9.1%) attributed to chronic obstructive lung disease in developing countries in Asia. It may also be partly responsible for the relatively high incidence of lung cancer in women in north China, many of whom spend much of their day in such confined and polluted spaces. Further research is needed to establish the relative importance of indoor pollution and the extent of its interaction with tobacco use and other factors.

The Interdependence of Health in All Countries

Considered sui generis, none of the facts I have described justifies my initial statement that the activities in each part of the world affect conditions everywhere else and that we need to examine what is happening in other countries when we examine the relationship between health and the environment in our own. I could have cited the spread of AIDS in Africa as an example, had I not specifically excluded the risk of AIDS from consideration by excluding from environmental factors those that individuals can largely choose to avoid. What I had in mind, however, was the much more complex effect of human activities on the world's climate and resources.

Climatic Change

Until recently any suggestion that human activity could have global effects would have been dismissed as science fiction. Now, however, two effects of human activity have been confidently predicted: an increase in the amount of shortwave ultraviolet radiation reaching the earth and an increase in global temperature.

The first effect is produced by the dispersal into the atmosphere of chlorofluorocarbons from aerosol sprays, plastic foam blowers, and refrigerants, and of halons from fire extinguishers. These gases, which have been produced only since the end of the Second World War, can persist in the atmosphere for up to 100 years, and their concentration in the atmosphere increased during the 1980s by about 5% per year. Chemically they interact with ozone; they are depleting the ozone layer over the mid-latitudes in the northern hemisphere by 4% to 5% per decade and the global ozone layer by 2% to 3% per decade.[17] Much greater depletion must be expected, even if the United Nations agreement signed by 93 countries in June 1990 is adhered to: namely, to stop producing chlorofluorocarbons by the end of the decade and to provide a fund of $240 million to help developing countries obtain alternatives.

The direct effect on health of this global change may be relatively small—an increase in the risk of melanomas and of basal and squamous cell carcinomas of the skin by perhaps 20% to 30%, which could be avoided by changes in personal behavior. Much greater effects may be produced indirectly through damage to phytoplankton in the oceans and other critical organisms in the food chain.

The second effect is more long-term, more potentially damaging, and much more difficult to prevent. It arises from the progressive accumulation of "greenhouse gases" in the atmosphere (principally carbon dioxide, methane, and nitrous oxide; carbon dioxide is responsible for half the effect) and is thought likely to increase mean global temperatures by 1.0°C within 30 years and 3.0°C by the end of the next century. Such a change would be almost as big as the fall of 4°C that occurred in the last ice age. A change of this order would be expected to reduce the cover of snow and ice in the Arctic and Antarctic and to cause a rise in sea level of around 20 cm in the next 40 years and around 65 cm by the end of the century, a general redistribution of weather patterns, and drastic local changes in the distribution of rain. There are many uncertainties that make precise prediction impossible. Again, the direct effect of these changes on health may be small, but the indirect effects could be substantial; such indirect effects might include the aggravation of food shortages, the extension of areas favorable to parasitic disease, and a vast increase in migration. Reversal of such effects would take many generations; the time to prevent them is now. Prevention would entail increasing the tree cover of the globe (to increase the removal of carbon dioxide) and reducing the combustion of fossil and biomass fuels, whereas the reality is that the reverse is certain to happen for many decades to come, due to the inexorable increase in the world population and the need for developing countries to increase their productivity if they are to raise their standards of health to that of the industrialized world.

Impact of Population Growth

The estimated world population in 1800 was about 1 billion, in 1900 about 2 billion, and in 1987 about 5 billion. The world population is still increasing at the rate of about 26 persons every 10 seconds. The increase, of course, is due to the fact that mortality rates are falling faster than fertility, and unless fertility falls much faster in the future than it now appears likely to do, the population is expected to reach about 8.2 billion in 2025 and perhaps 10 to 12 billion before it eventually stabilizes.

Whether agricultural production can increase sufficiently to provide a reasonable diet for so many people is seriously in doubt. Up to now better agricultural practices, the greater use of fertilizers, and the introduction of new varieties have enabled standards to rise in many countries despite an increasing population, but the productive capacity of the earth is limited. Myers[18] has calculated that even if we all ate a vegetarian diet and shared our food equally the biosphere could not support more than about 6 billion people; if everyone were to require 35% of their calories to come from animal products, as is now the case in North America, the biosphere could adequately support only about half its present population.

In practice, the biosphere is not being farmed ideally. Not enough food is distributed from areas where it is overproduced to areas where it is needed. Where industrial production has not developed to a level that enables the country to buy food, malnutrition is widespread, and the overuse of marginal soils in some of the poorest areas is causing progressive desertification, with a consequent reduction in the potential for food production in the future.

By far the largest increase in population is taking place in the developing countries, and this is likely to remain true for many years to come. Table 9 shows the population of the industrialized and developing countries in 1950, the percentage increases in both groups between 1950 and 1985, and the expected increases by 2025. Both the number and the proportion of people who will require an increased standard of living, if their standard of health is to be raised to a level that would be regarded as acceptable in our own countries, is increasing steadily. If past experience is anything to go by, such an

TABLE 9—Predicted Growth of World Population, 1990 to 2025

Type of Country	1990 Population (millions)	Percentage change since 1990	
		1995	2025
Industrialized	832	+41	+68
Developing	1684	+118	+304

Note: Adapted from the Report of the WHO Commission on Health and the Environment, Geneva: World Health Organization, 1992.

improvement cannot be achieved without either dictatorial imposition of behavior or a greatly increased use of biomass and fossil fuels, with a corresponding increase in the production of carbon dioxide. Moreover, the increasing population will require more land for agricultural purposes and more wood for fuel. Consequently, the rate of destruction of tropical rainforests—which are at present being reduced by about a hectare every 20 seconds—will be accelerated, thus reducing the ability of the biosphere to remove carbon dioxide and speeding further the processes that lead to uncontrollable climatic change.

Conclusion

The environmental factors that I have reviewed form only one component of the many factors that are shaping the health of the peoples of the world, and I do not wish to give the impression that others are unimportant. On the contrary, the outlook for health in Great Britain and the United States in the immediate future is highly encouraging (apart from the increasing risk of AIDS) principally because of the possibility of reducing the risk of premature death in middle age by reducing the prevalence of smoking, modifying diet, and improving treatment. In some developing countries one of the urgent tasks is to prevent the spread of the epidemic of smoking that the tobacco industry is striving to cause. If even the present smoking habits of young people are continued into later life, tobacco can be expected to cause 10 million deaths a year worldwide by the 2020s.[19]

Environmental factors, nevertheless, continue to be immensely important in both industrial and developing countries—although not primarily the factors that are often singled out for special emphasis. That modern industry has given rise to many occupational hazards and to the dispersal of small amounts of many potentially toxic chemicals is not open to question. In Great Britain and the United States, however, both are now tightly controlled, and public exposure to them has been so small that any associated risk is difficult if not impossible to detect. The major environmental hazards that now need to be tackled are quite different: poverty, the pressure of population growth, and the production of greenhouse gases.

Poverty is immensely difficult to deal with and I know no easy solution. In the United States a beginning could perhaps be made, on a pilot basis, by redistributing the $4 billion currently spent each year by Superfund on cleaning up waste sites that, according to the National Research Council's committee,[10] have been selected on the basis of engineering feasibility rather than on degree of hazard. The money could be used to improve conditions in selected poverty-stricken areas. To make effective use of the funds, however, according to WHO's Commission on Environment and Health,[20] it would be essential to involve the local population in decisions about their use.

In developing countries the greatest contribution that biomedical science can make is to help reduce the pressure of population growth. Fertility rates fall as economic and educational standards rise and mortality is reduced, but the world, and developing countries in particular, cannot wait for this to happen. The demands placed on resources by the growing population and growing consumption are too severe. Wealthy countries, for their part, can contribute to worldwide population reduction in two ways. First, they can reduce their fertility rates to reproduction level or below (as has already been done in the United States and Great Britain), not only to set an example, but also because their citizens make the biggest demands on limited resources. The citizens of the industrialized nations consume about 10 times as much fossil fuel as do the citizens of developing countries, with US citizens consuming 40 to 50 times as much as the citizens of the poorer African and Asian countries. Second, the wealthy countries can provide funds for the support of family planning worldwide and for research into th development of contraceptives (both male and female) that require a minimum of medical intervention.

Lastly, and equally difficult, is the adoption of policies to reduce the production of greenhouse gases. The responsibility for this reduction is squarely on the shoulders of the countries that are already industrialized, for the developing countries cannot be expected to refrain from increasing their production at least to the level required to enable them to have the standard of health to which we are accustomed in North America and Europe. No one method will secure the necessary result, but it can be achieved by the combination of several methods. We must achieve greater efficiency in the use of energy by insulating to avoid heat loss and by substituting public for private transport, and we must decrease our reliance on fossil fuels, possibly by making greater use of nuclear power and certainly by developing methods for obtaining energy from renewable sources (winds, tides, and sun).

The achievement of these aims is essentially a political problem, but that is no reason for us not to be concerned with it. By increasing our life spans and raising our standard of living, medicine and industrial development have indirectly created the problems resulting from our higher consumption of natural resources, and we cannot escape partial responsibility for them. □

Acknowledgments

This paper was presented as the first Stallones Memorial Lecture at the University of Texas School of Public Health at Houston, January 23, 1992.

The author is grateful to Dr. C. Hennekens and Dr. F. Speizer for advice about some aspects of US vital statistics and to Sir Crispin Tickell for advice about food supplies and climate change.

References

1. Lopez AD. *Causes of death in the industrialized and the developing countries: estimates for 1985.* Washington, DC: The World Bank, Population and Human Resources Department; 1991. Health Sector Priorities Review 20.
2. Otten M, Teutsch SM, Williamson DF, Marks JS. The effect of known risk factors on the excess mortality of black adults in the United States. *JAMA.* 1990;263:845-850.
3. McCord, Freeman. Excess mortality in Harlem. *N Engl J Med.* 1990;322:173-177.
4. Council on Ethical and Judicial Affairs. Black-white disparities in health care. *JAMA.* 1990;263:2344-2346.
5. Ellison J McK, Waller RE. A review of sulphur dioxide and particulate matter as air pollutants with particular reference to effects on health in the United Kingdom. *Environ Res.* 1978;16:302-325.
6. Waller RE. The effects of sulphur dioxide and related air pollutants on health. In: *International Symposium on Sulphur and the Environment.* London: Society of Chemical Industry; 1979:171-177.
7. Pike MC, Gordon RJ, Henderson BE, Menck HR, Soohoo J. Air pollution. In: Fraumeni JF, ed., *Persons at High Risk of*

Cancer: An Approach to Cancer Etiology and Control. New York: Academic Press; 1975:225–238.
8. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst.* 1981;66:1191–1308.
9. Ames BN, Gold LS. Chemical carcinogenesis: too many rodent carcinogens. *Proc Natl Acad Sci U S A.* 1990;87:7772–7776.
10. National Research Council. *Environmental Epidemiology.* Vol. 1. *Public Health and Hazardous Wastes.* Washington, DC: National Academy Press; 1991.
11. Peto R, Lopez AD. Worldwide mortality from current smoking patterns: WHO consultative group on statistical aspects of tobacco-related mortality. In: Durston B, Jamrozik K, eds. *Tobacco and Health 1990: The Global War.* Perth, Western Australia: Proceedings of the 7th World Conference on Tobacco and Health, April 1990.
12. Doll R. Progress against cancer: an epidemiological assessment. *Am J Epidemiol.* In press.
13. Doll R. Are we winning the war against cancer? In press.
14. Knox EG, Lancashire RJ. *Epidemiology of congenital malformations.* London: HMSO; 1991.
15. International Clearinghouse for Birth Defects Monitoring Systems. *Congenital malformations worldwide.* Amsterdam: Elsevier Science Publishers; 1991.
16. World Health Organization. *World Health Statistics Annual, 1991.* Geneva: World Health Organization; 1991.
17. Kerr RA. Ozone destruction worsens. *Science.* 1991;252:204.
18. Myers N. *Mass extinctions: Palaeogeography, palaeoclimatology, palaeoecology.* Amsterdam: Elsevier Science Publications; 1990.
19. Peto R, Lopez AD. Tobacco-attributable mortality: global estimates and projections. In: *Tobacco Alert.* Geneva: World Health Organization; January 1991:4–5, 7.
20. Commission on Health and the Environment. *Our planet, our health.* Geneva: World Health Organization; 1991.

Commentary: Environmental Disease—A Preventable Epidemic

ABSTRACT

Toxic environmental diseases are highly preventable causes of morbidity and mortality. Toxic diseases in the work environment cause an estimated 50 000 to 70 000 deaths and 350 000 new cases of illness each year in the United States; the asbestos pandemic will ultimately claim at least 300 000 lives; pediatric lead poisoning is epidemic, and an estimated 3 to 4 million US preschool children have blood lead levels above 10 μg/dl and could suffer long-term neuropsychological impairment. Prevention of environmental diseases can be achieved through legislation and regulation that control common-source exposures to chemical toxins. Modification of personal behaviors, such as tobacco and alcohol consumption, complements but does not replace control of toxic environmental exposures. (*Am J Public Health.* 1992;82: 941–943)

Philip J. Landrigan, MD, MSc

Introduction

Disease caused by toxic chemicals in the environment is a substantial yet highly preventable cause of morbidity and mortality in the United States and around the world. Consider:

• *Lead poisoning.* Lead poisoning is epidemic among young US children. The Centers for Disease Control (CDC) estimates that 3 to 4 million American preschool children have blood lead levels above 10 μg/dl.[1] Blood lead levels in this range in young children have been shown to cause depression of neurological and psychological function,[2–4] effects that appear to be permanent.

• *Asbestos.* Asbestos in the work environment has created an absolute disaster.[5] Although the hazards of asbestos were known in the 1930s, billions of tons of asbestos were used in the United States from then until the 1980s. Millions of workers were exposed, and tens of thousands of homes, schools, and other public buildings were contaminated.[6] By the year 2000, an estimated 300 000 American workers will have died of the diseases caused by asbestos: principally, asbestosis, lung cancer, and malignant mesothelioma. It is anticipated that deaths will continue to occur well into the 21st century.[7] With the overseas spread of asbestos-containing building materials, the pandemic is now extending to the Third World.

• *Occupational disease.* Studies conducted in New York State have estimated that 50 000 to 70 000 US workers die each year of chronic occupational disease resulting from past exposures to toxic substances.[8] These diseases encompass a broad range: besides the lung cancer and mesothelioma of asbestos workers,[5] they include cancer of the bladder in dye workers[9]; leukemia and lymphoma in workers exposed to benzene[10] and ionizing radiation[11]; chronic bronchitis in workers exposed to dusts[12]; disorders of the nervous system (including possibly dementia, Parkinson's disease, and motoneuron disease) in workers exposed to pesticides, solvents, and certain other neurotoxins[13]; renal failure in workers exposed to lead[14]; and cardiovascular disease in workers exposed to carbon monoxide[15] and carbon disulfide.[16]

A major impediment to assessing the contribution of synthetic chemicals to the causation of cancer and other workplace diseases is the fact that only 20% of the

Philip J. Landrigan is professor and chair in the Department of Community Medicine, Mount Sinai School of Medicine, New York, NY.

Requests for reprints should be sent to Philip J. Landrigan, MD, MSc, Department of Community Medicine, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10009.

The Potential Impact of Epidemiology on the Prevention of Occupational Disease

David H. Wegman, MD, MS

ABSTRACT

This presentation reviews occupational epidemiology as a foundation for workplace disease prevention activities. By examining descriptive, etiologic and intervention occupational epidemiology studies, a range of opportunities are illustrated where epidemiology has played, or could play a principal role in guiding preventive efforts. Descriptive studies presented include ones based on vital records, on epidemic investigations, cross-sectional surveys, and surveillance. Etiologic studies review the largely successful development of knowledge for lung cancer and asbestos exposure for pulmonary effects of isocyanate exposures. However, attention is also directed to the need for etiologic studies of work environment risks for both cardiovascular and musculoskeletal disease. Finally importance is placed on the too infrequent epidemiologic studies of intervention. Historical examples of control of large risks from nickel cancers and silicosis are balanced with more recent examples of successes at reducing smaller risks of cardiovascular disease and oil acne. Throughout, emphasis is placed on the importance of reintegrating the academic discipline of epidemiology into the application of study findings to prevention of workplace risks. (*Am J Public Health.* 1992;82:944–954)

Introduction

Over the first half of this century public health professionals used the tools of epidemiology primarily for the prevention of infectious disease. The appropriate knowledge was gained in the standard public health curriculum and communication between academic and practicing epidemiologists was relatively easy. Much of the epidemiologic research of that time was carried out by professionals who served as or closely collaborated with public health epidemiologists in studying both the cause and the prevention of epidemics.

More recently, however, as attention has shifted to the study and prevention of noninfectious disease, a gap has appeared between the contemporary discipline of epidemiology and the practice of public health. Szklo[1] suggested that this gap relates to a generational difference in epidemiologists, "a difference expressed by the fact that those who produce data needed for preventive programs (namely epidemiologists) and those in charge of carrying out such programs very often are not, unlike in the good old days, the same people at all."

These comments reflect the fact that epidemiologic study in the United States has come to focus most often on well-planned, large, time-consuming studies of chronic disease. Such studies have benefited from the development of new methodology but also have required less direct interaction with public health agency personnel and use of public records. As a result, as the tools of epidemiology have become more powerful and the scientific base of the discipline has broadened, epidemiologists have turned more and more to the excitement associated with seeking new knowledge and left to others the more mundane task of implementing that knowledge.

The evolution of epidemiologic methods in studying noninfectious disease has been of particular benefit to the study of occupational disease and has resulted in expanded knowledge about health risks in the work environment. Consequently, many bright young scientists have been attracted into the discipline. The positive developments in epidemiology as an academic discipline, however, have been accompanied by increasing distance between the development of and the practical application of new knowledge. Consequently, this commentary is directed at rebuilding bridges between academia and public health agencies, over distances that, I believe, are short and through terrain that is exciting.

In exploring applications of epidemiology to the prevention of occupational disease, it is useful to examine potential opportunities according to the recognized study categories: descriptive, etiologic, and intervention. The first two of these are commonly considered, but there is an unusual opportunity to emphasize the third when undertaking occupational epidemiology studies.[2] *Descriptive studies* provide information for setting priorities, identifying hazards, and formulating hypotheses for new occupational risk. It is surprising how often such studies have

The author is with the Department of Work Environment, College of Engineering, University of Massachusetts Lowell, Lowell, Mass.

Requests for reprints should be sent to David H. Wegman, MD, MS, Department of Work Environment, College of Engineering, University of Massachusetts Lowell, Lowell, MA 01854.

This paper was submitted to the Journal December 3, 1991, and accepted with revisions March 24, 1992.

proved sufficient to initiate intervention. *Etiologic studies* are planned examinations of causality and the natural history of disease. These studies have required increasingly sophisticated analytic methods as the importance of low-level exposures are explored and greater refinement in exposure-effect relationships is sought. Finally, *intervention studies* seek to demonstrate that steps taken to reduce work-related risk factors do, in fact, result in measurable disease reduction or improvement in the health of the work force.

Much of the emphasis in modern epidemiology has been on studies of etiology. These, of course, need to remain prominent, but not to the exclusion of the other essential roles for epidemiology. The following examination of accomplishments in each of the three areas, therefore, is designed to kindle an enthusiasm and sense of activism for the full range of opportunities to use epidemiology in the prevention of occupational disease.

Descriptive Epidemiology

William Farr offered one of the earliest systematic applications of epidemiologic descriptive analyses to occupational risks. He took advantage of Great Britain's vital records reports on occupation-related mortality to illustrate the striking mortality excesses among metal miners.[3] Using the data shown in Table 1, he demonstrated that the metal miners of Cornwall experienced higher mortality from all causes than did the general population. To examine whether the excesses were in some way due simply to underground work, he compared metal miners with coal miners. He concluded that mining alone did not provide a sufficient explanation of the excess risk.

After examining cause-specific mortality Farr was able to show that the excess was concentrated in deaths from pulmonary disease. Table 2 illustrates the results of his efforts to examine whether the excess was unique to Cornish metal miners. Pulmonary disease mortality in lead miners from two other parts of Great Britain (Northern counties and Wales) was compared with that of the Cornish miners. Despite some regional differences, it is reasonable to conclude that the excesses in the Cornish and other miner cohorts were due to the metal mining itself. An interesting historical note is that, at this early date, Farr introduced healthy worker selection as an explanation for the pulmonary disease mortality findings among the youngest miners in the two comparison regions.

TABLE 1—Average Annual Deaths per 1000 from All Causes: Cornwall Metal Miners Compared with All Other Cornwall Deaths and Northern Coal Miners

Age Range	Cornwall Metal Miners Mortality Rates (1860–62)	Rate Ratios: Metal Miners/General Population	Rate Ratios: Metal Miners/Coal Miners
15–25	8.90	1.25	1.06
25–35	8.96	1.01	1.06
35–45	14.30	1.43	1.41
45–55	33.51	2.27	1.99
55–65	63.17	2.63	2.58
65–75	111.23	1.89	1.71

Note. Adapted from Farr.[3]

TABLE 2—Ratio of Pulmonary Disease Mortality Rates in Metal Miners Compared with All Other Residents in Three Regions

Age Range	Cornwall Metal Miners	Welsh Lead Miners	Northern Lead Miners
15–25	1.14	0.89	0.88
25–35	1.06	0.72	1.24
35–45	1.86	1.96	3.34
45–55	4.55	2.08	4.45
55–65	8.34	2.89	5.75
65–75	4.30	2.85	3.06

Note. Adapted from Farr.[3]

The high prevalence of employment in one major type of work and the high prevalence of disease due to that work provided Farr a rich soil to till. Data of this type are not as likely to provide such dramatic findings today. However, the existing decennial publications from Great Britain and the occasional reports from other geographic locations provide a valuable resource on mortality experience for different occupational groups that, as yet, are incompletely explored for hypotheses about work-related risks.[4]

Descriptive studies of illness outbreaks can also be seen to play a role in prevention. This fact is illustrated by an outbreak investigation that was initiated by a report from a hospital clinic. Eleven employees from a nearby polyurethane production plant visited the clinic, all with symptoms of urinary retention. A review of the medical records of these workers found no common diagnosis and so a survey of the entire work force (208 employees) was planned.

An appropriate case for study was arbitrarily defined as the report of at least two of four specific urinary tract symptoms. Sixty-three percent of the production workers met the case definition. A plot of onset times (Figure 1) showed strikingly how the outbreak coincided with the introduction of a new chemical catalyst. This observation formed the basis of a recommendation that the catalyst be replaced. It can be seen that the outbreak abruptly ceased after the catalyst was removed. This and a parallel epidemic investigation in another city led the producer to remove the catalyst from commercial use. Publication of the findings[5] also led to a collaboration with toxicologists who are now studying the biochemical mechanism of the toxic action and also exploring the potential neurotoxicity of related compounds.[6–8]

Similar occupational disease outbreak investigations provide lessons about the growth of knowledge about occupational disease and the opportunities for its prevention. Other sources of descriptive knowledge are also documented, including reports based on such simple but elegant insights as relating disease in the community to the correct denominators. For example, Aksoy reported an evaluation that was based on 26 acute leukemia

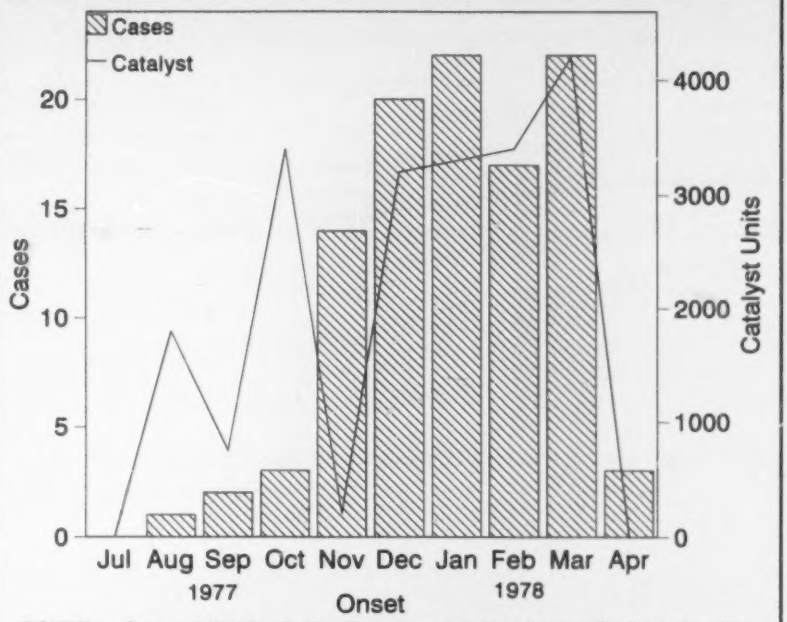


FIGURE 1—Cases of bladder dysfunction among workers by month of onset and by amount of catalyst used. Adapted with permission from Kreiss et al.,[5] Copyright © 1980, American Medical Association.

TABLE 3—Percentage of Workplaces in which Environmental Hazards Were Found, by Size of Workplaces (Thailand Workplace Survey)

Hazard Type	Workplace Size: Less than 50	Workplace Size: 51–150
Lead fumes/vapors	20	15
Airborne irritants	46	52
Heat stress	30	33
Excessive noise	50	20
Inadequate lighting	54	30
Accident potential	66	50

Note. Adapted from El Batawi and Husbumrer.[11]

cases that were among those newly diagnosed in Istanbul over a 7-year period. He observed that these 26 cases occurred in shoe industry workers, whereupon he used census data to estimate that a total of 28 500 workers were employed in Istanbul's shoe industry. Using a quasi–person-years approach, he then calculated a 7-year crude incidence rate for leukemia of 13 per 100 000 among shoe industry workers, which could then be compared to the estimated 6 per 100 000 in the general population of Istanbul.[9] Aksoy hypothesized that the difference in risk was due to exposure to benzene contained in the glues used in the shoe industry, which included many small workshops and few exposure controls.

Although this descriptive analysis has been criticized for a number of limitations, the hypothesis of benzene-related leukemia has since been confirmed. Furthermore, at least two important public health consequences can be traced to this simple descriptive study. First, the hypothesis presented in this study served to stimulate several well-planned etiologic studies, results from which were used to guide reduction in established exposure limits.[10] Second, an increase was seen in the pace of voluntary substitution or control of benzene even as the hypothesis was being tested.

Well-planned, descriptive field surveys can also provide important guidance for targeting more general prevention activities. In Thailand, a stratified sample of different types of small and medium manufacturing firms identified a need for a variety of occupational health services.[11] In addition to the wide distribution of hazards shown in Table 3, the survey provided evidence that over a third of printing and battery workers had some symptoms or signs of lead poisoning and a disproportionate number (almost 50%) of textile workers suffered from upper respiratory infections. The survey led the Health Ministry to direct new attention to workers' health with development of educational programs for workers and professionals, establishment of a new occupational hygiene laboratory, and passage of legislation to require more health care for small industries.

Finally, another form of descriptive epidemiology worthy of mention is surveillance. Surveillance continues to play a central role in the identification and control of communicable diseases. It has not so far proved as valuable when applied to occupational disease, although as far back as William Farr there are examples of moderately successful forms of surveillance.

An example is provided from the annual US occupational injury reports. Generally these reports have attracted little attention, but there is intense interest in determining the cause of the recent striking rise in reports of repetitive trauma disorders in the United States.[12] Between 1972 and 1982 the number of reports of such disorders was close to 20 000 per year. Since 1983, however, there has been a steady rise in reports, and the number is now over three times higher. Initial examination of these trends suggests that they cannot be explained by increased reporting alone but are likely to represent a real increase in disease incidence as well.

New approaches to occupational disease surveillance are being explored.[13] But in the spirit of a broader concept of occupational epidemiology, I would like to call attention to a less familiar form of surveillance, hazard surveillance. Hazard surveillance is directed at work exposures rather than disease, but is consistent with MacMahon and Pugh's definition that epidemiology is the study of the "distribution and determinants of disease."[14] A major benefit of hazard surveillance is that it should provide more timely information on risks related to agents that cause chronic long latency disease such as mesothelioma.[15] In contrast to the mesothelioma registry, hazard surveillance of asbestos exposures could show trends predictive of future disease without waiting to count the actual cases. Surveillance of exposure trends can be seen as following trends in disease *risk,* rather than in disease itself.

The potential for hazard surveillance can be illustrated with information from the US Occupational Safety and Health Administration (OSHA) compliance activity. A detailed analysis of lead exposures collected over the decade following the introduction of a new lead standard

(Table 4) shows that problems persist in the six major lead industries that were the focus of the standard.[16] Not shown is the fact that, despite issuance of the new standard in 1978, there is no trend toward lower exposures.

Detailed review of job-specific exposures called attention to several specific jobs in which overexposure was most serious. A particularly important and unexpected finding emerged from the analysis of exposures *outside* the six major industries. Excessive lead exposures in the job title "painter" appear in a number of different industry groups. Although structural painters have long been known to risk lead exposure, these data showed that small numbers of painters in a wide variety of industries commonly had elevated lead exposures. Among the 848 inspections of plants that were not part of a lead industry, one third of the test samples exceeded the exposure limit, 39% involving this one job title. These results from epidemiologic surveillance of hazards characterized the need for a program targeted to this hidden generic problem.

TABLE 4—Inspections in which Observed Levels Exceeded Permissible Exposure Levels (PELs) and Samples in which Levels Exceeded Permissible Levels: OSHA Inspection Data, 1979–1985

Industry	No. (%) of Inspections with Median Levels > PEL	No. (%) of Samples with Levels > PEL
Inorganic pigments	25 (52)	143 (60)
Primary lead	5 (40)	72 (83)
Secondary nonferrous metals	86 (43)	718 (49)
Brass/bronze/copper foundaries	118 (42)	822 (44)
Storage batteries	87 (57)	969 (56)
Primary batteries (dry and wet)	18 (33)	180 (38)

Note. Adapted with permission from Froines et al.,[16] Copyright © 1990, Wiley-Liss.

Etiologic Epidemiology

Doll has suggested that to indicate etiology, epidemiologic studies should meet at least four requirements. They should show (1) positive associations between exposure and disease in populations with known exposure; (2) associations shown not to be due to bias, confounding, or chance; (3) associations that are replicated in different circumstances; and (4) associations that vary appropriately with dose and period of exposure.[17] He argues, however, that the International Agency for Research on Cancer requirement for the association also to be strong is inappropriate for public health because that requirement is likely to emphasize only large but rare risks. In the context of prevention, epidemiologic studies of etiology are most useful when they (1) improve understanding of exposure and exposure-response relationships, which provides information essential for primary prevention; and (2) advance knowledge about host factors and early markers of effects, which contributes to efforts at secondary prevention. These features will be illustrated with examples from the study of occupational cancer and from the study of occupational pulmonary disease.

The evolution of our current understanding of asbestos-related cancer illustrates the importance of epidemiologic study in developing and examining hypotheses about the nature of the risk. During the first half of the 20th century case reports and descriptive studies provided suggestive evidence that asbestos was a carcinogen. However, the first well-planned epidemiologic studies to address the question were not carried out until the mid-1950s. Then in 1954 and 1955 the suspicion that asbestos was carcinogenic, not just fibrogenic, became a specific hypothesis after the almost simultaneous publication of Doll's cohort study of asbestos textile workers[18] and Breslow's case-control study of lung cancer.[19] Doll's study documented an 11-fold excess risk in textile workers, while Breslow's study provided evidence that asbestos-related lung cancer mortality was not due to cigarette smoking alone.

The following quarter-century saw probably the most intensive effort at epidemiologic investigation directed at any work environment risk to date. When such efforts were summarized by McDonald in 1987, a total of 50 cohort studies of men and women had been carried out.[20] A number of different findings useful for prevention emerged from these studies. The characterization of cumulative exposure and lung cancer risk has now been provided for several different industry groups. For example, Dement showed convincing evidence of the high risk associated with chrysotile exposures in textile work,[21] and McDonald provided equally convincing evidence that workers exposed during mining of the same material were at much lower risk.[22] McDonald has illustrated this (Figure 2) by applying uniform calculations of relative risk to several different cohorts.[23]

A separate finding from several other studies has been that the components of cumulative exposure (duration and intensity) have been shown to contribute independently to risk. This was demonstrated even when intensity could be categorized only as light, moderate, and severe,[24] as well as when average dust categories could be estimated.[25] A third lesson from studies of asbestos and lung cancer derives from one of the strengths of large cohort studies, the opportunity to examine small risk differences. In several asbestos cohorts it has been possible to document the effects of asbestos independent of cigarette smoking and in at least one the risk was shown to be more than additive.[26]

Thus, epidemiology studies have contributed importantly to providing answers to a great number of questions about asbestos risk. Severe restrictions have now been adopted to control use or exposures in most developed nations. Still, interesting questions remain. Some, such as why different fiber types are associated with different magnitudes of risk, will probably benefit more from experimental studies than from further epidemiologic study. Others, however, are probably well suited to an epidemiologic approach—for example, whether exposure level alone or other cofactors are the most likely explanation for the striking risk differences for chrysotile when exposure occurs in textile factories in contrast to exposure in mining, milling, or the manufacture of friction products.

The epidemiologic study of an occupational cancer risk provides important evidence to guide primary prevention activity. However, when it comes to cancer there are few tools available to the epidemiologist that allow much of a contribution to knowledge useful for secondary prevention. This is not the case for respiratory disease, as can be illustrated by the contribution of epidemiology to studying isocyanate-related acute pulmonary disease.

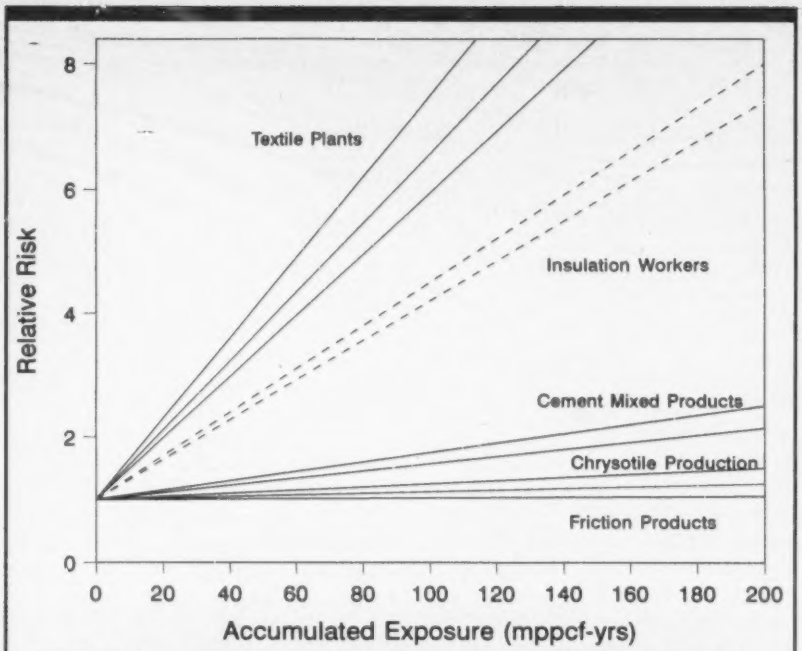


FIGURE 2—Lung cancer risks for 10 different occupational cohorts by estimated cumulative exposure to asbestos fibers. Adapted with permission from McDonald,[23] Copyright © 1984, *Annals of the Academy of Medicine, Singapore.*

TABLE 5—Cross-Shift Change in 1-Second Forced Expiratory Volumes (FEV_1), by Smoking History, for Workers Exposed to Low Levels of Toluene Diisocyanate

	FEV_1 (liters)		
	Preshift	Postshift	Cross-Shift Change*
Total Group	3.79	3.57	0.22
Never smokers	3.73	3.56	0.17
Ever smokers	3.84	3.60	0.24

Note. Adapted from Peters et al.[28]
*All values are significant by Student's *t* test at $P < .01$.

TABLE 6—Four-Year Change in 1-Second Forced Expiratory Volumes (FEV_1), by Smoking History and by Level of Exposure to Toluene Diisocyanate

Average Daily Exposure (ppb)	Smoking History		
	Never	Ex-Smoker	Current
< 2.0	0.013+	0.006	−0.014
2.0–3.5	0.005	0.004	−0.030
> 3.5	−0.084	−0.044	−0.047

Note. Values are given in liters per year. Table adapted from Wegman et al.[32]

In 1951, case reports of isocyanate-related pulmonary effects called attention to a new cause of occupational asthma.[27] Early epidemiologic studies were not able to add much to the case reports because those with asthma often left the work environment. Then, in the late 1960s, Peters and his colleagues initiated studies of workers exposed to very low levels of isocyanates. In very low exposures, few cases of asthma and thus little healthy worker selection bias were anticipated. Moreover, Peters et al. took advantage of the ease of field measurement of pulmonary function to test workers frequently over a work week both before and after their work shifts. In this way they were able to identify evidence for short-lived reactions to environmental exposures.

Significant changes in 1-second forced expiratory volumes (FEV_1) over the work shift were demonstrated (Table 5), as well as the fact that acute changes were more striking in those who had ever smoked cigarettes.[28] A 3-year prospective study of the same workers demonstrated that the losses were not just acute; exposed workers experienced accelerated losses over 3 years and there appeared to be a correlation between the magnitude of cross–work shift change and the accelerated functional loss.[29,30] The findings suggested that isocyanates were not only asthmagens but also caused nonspecific bronchoconstriction. It was further suggested that those most likely to suffer accelerated loss in lung function could be identified by cross-shift measurements of pulmonary function. Further studies supplemented these findings with evidence for the relationship of chronic function losses to exposure level (Table 6).[31,32]

Peters et al.'s work was elegant in its simplicity and clarity. It also provided evidence that cross-shift change in pulmonary function could be used both as a measure of outcome for epidemiologic studies and, potentially, for medical monitoring as well. Peters et al.'s study may have underestimated the importance of isocyanates because it was, as is often the case, a study of current employees—that is, a survivor population. Weill and his colleagues had an unusual opportunity to undertake a prospective study of employees who were exposed for the first time in a new production process. He used an approach similar to that used by Peters et al. and was able to document dose-related effects of chronic isocyanate exposure while directly accounting for healthy worker selection.[33,34]

These epidemiologic studies provided sufficient evidence that isocyanates were not only a problem for a small population of sensitized workers, but that the agents presented some risk for all exposed workers. Control efforts, therefore, could not simply be directed at the protection of a small subset of those exposed. The studies also suggest the value of the cross-shift measurement of pulmonary function and the assessment of accelerated loss of function over time as effective measures for monitoring and surveillance of exposed

populations. These add importantly to the tools necessary for both primary and secondary prevention efforts.

It is not accidental that these two examples of etiologic epidemiology come from studies of cancer and respiratory disease; these diseases have been given the greatest attention by epidemiologists. It is worth briefly examining two other very common disease groups that have not received adequate study by occupational epidemiologists: cardiovascular disease and musculoskeletal disease. If important work-related risk factors were well characterized for either of these disease groups, such factors would likely be more susceptible to preventive interventions than are the personal risk factors now known.

Cardiovascular disease is recognized as the leading cause of death in industrialized nations and is equally important in many developing nations. Through epidemiologic studies of general populations, a number of risk factors useful for prevention have been identified. A recent review calls attention, however, to important nonchemical risk factors for cardiovascular disease and concludes that there are several attractive hypotheses needing immediate attention by occupational epidemiologic study.[35]

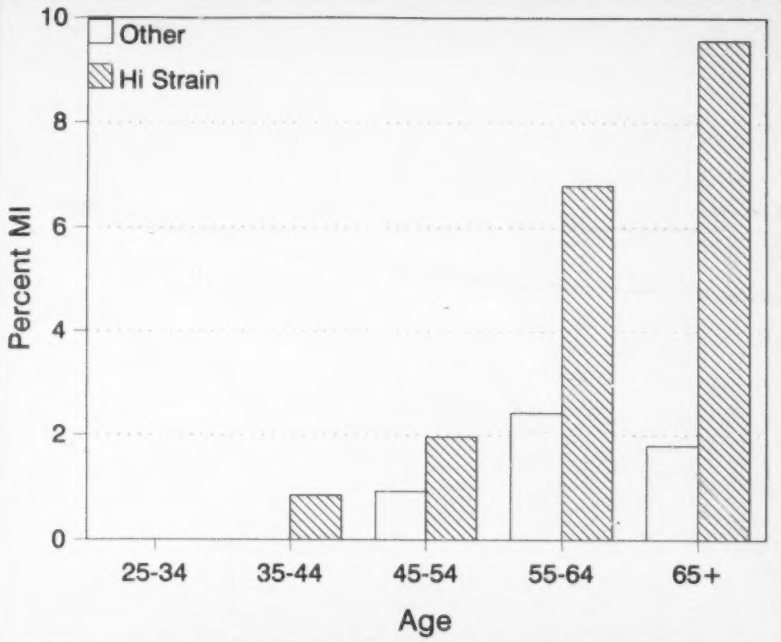


FIGURE 3—Prevalence of myocardial infarction by age and job strain (high job strain applies to those jobs judged to have high psychological demand and low worker control) among employed males surveyed in Cycle 1 of the US National Health and Nutrition Examination Survey (NHANES). Adapted with permission from Karasek et al.,[36] Copyright © 1988, American Public Health Association.

One of the possibly most important risk factors (or complex of factors) concerns the organization of work. A leading idea used in examining this risk factor is Karasek's concept of work strain. Karasek et al. have successfully captured work strain with a two-dimensional job rating scheme (psychological demands on the worker and control over those demands). An illustration (Figure 3) of Karasek et al.'s work on high strain is seen in an analysis of data from one of the US National Health and Nutrition Examination Surveys.[36] Kristensen reviewed several Scandinavian epidemiologic studies, using the two dimensions to define high strain, and concluded that there is convincing epidemiologic evidence that cardiovascular disease is associated with jobs in which there is high psychological demand matched with low worker control of the demand.[37] This effect has been noted in the Framingham Heart Disease study as well.[38] Epidemiologic studies have also begun to examine other characteristics of work organization that may be equally important, such as job instability (e.g., Siegrist[39]) and shift work (e.g., Knutsson[40]).

A second inadequately studied risk factor is the role of physical inactivity present in the work environment. This factor is particularly important because it appears that the evolution of many jobs is toward increasingly more sedentary work. In this regard, Paffenbarger and his colleagues have contributed considerably to measuring the importance of inactivity. In one of their studies of longshoremen, Paffenbarger et al. noted that those whose jobs called for high physical activity showed low risk of sudden death.[41] The total energy expenditure of the longshoremen during the workday was of the order required by other physically demanding jobs, such as a tire builder in the manufacture of rubber tires. What was especially interesting, however, was the pattern of work that appeared to be protective against sudden death. Paffenbarger et al.'s work suggested that it was not just the level of activity but the intermittent peak energy expenditure that was protective. Further epidemiologic study of this finding may be instructive for designing jobs that require intense physical exertion so that the jobs provide the maximum preventive health benefit.

If epidemiologic studies of work organization and heart disease identify features of work that can be altered, the public health impact could indeed be great. Equally important, business leaders are likely to be willing to act on such findings. The payoff of such cardiovascular disease research could be very big. If the evidence is convincing, work organization can be redesigned and proper work schedules designed. In the United States and throughout the world, increasingly more attention is being directed at trying to contain the increase in costs of medical care. Throughout the world the high prevalence of cardiovascular disease and its related disability demands public health action.

It is recognized that the problems of disability associated with musculoskeletal diseases are also experienced in a wide variety of jobs, both commercial and industrial, physically demanding and largely sedentary, male- and female-dominated. One interesting research direction is the recent work in Finland directed at determining age-related musculoskeletal *capacity*. Work of this type could lead to critical data not only on the role of physically strenuous work in causing disability but also on the need to develop physiologic criteria for retirement age that address anticipated losses in musculoskeletal capacity.[42,43]

Despite increasing attention, further progress in the epidemiologic study of musculoskeletal disease has been limited

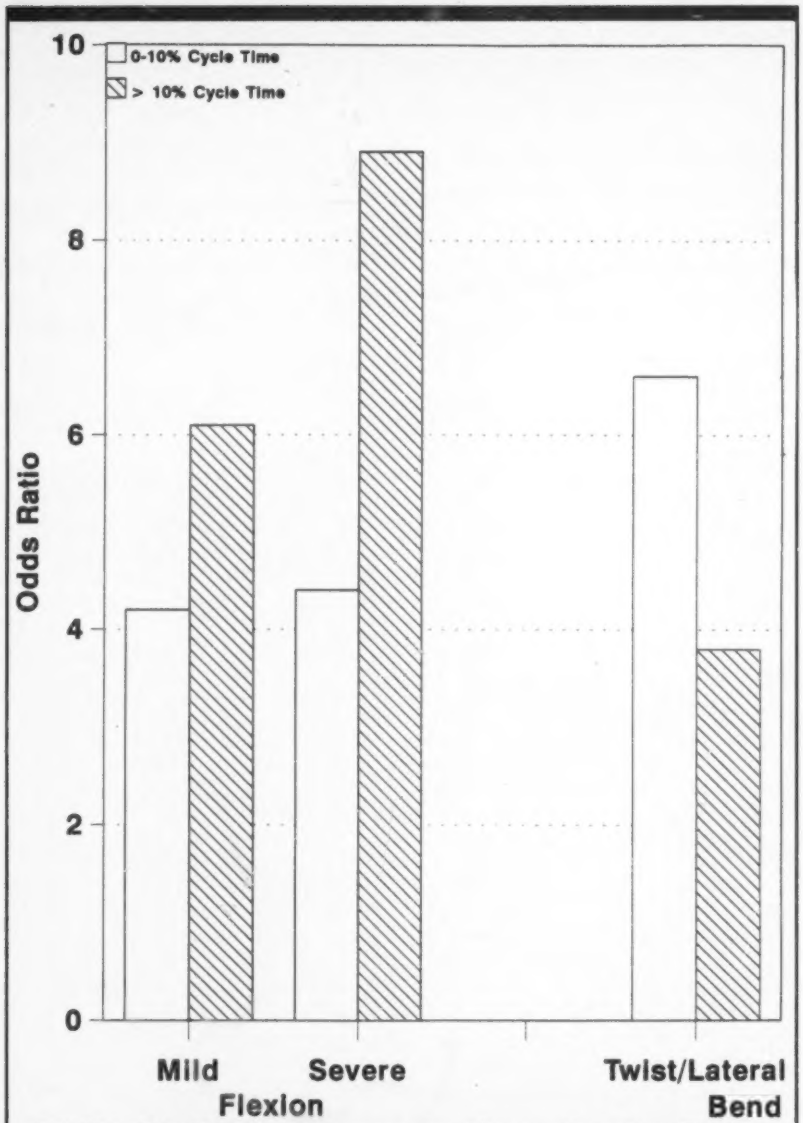


FIGURE 4—Risk of back disorders among automobile assembly workers with mild or severe use of forward flexion or with twisting and lateral bending; cycle time = time required to perform one complete set of production operations. Adapted from Punnett et al.[46]

by significant problems with proper measurement of both exposure and outcome. Recent reviews indicate that musculoskeletal diseases such as carpal tunnel syndrome[44] and different types of shoulder-neck syndrome[45] are associated with a variety of different job types. Existing knowledge of job risk factors, however, does not provide common features of these different jobs. An important priority, therefore, is to develop the means to identify and quantify relevant common microergonomic features of jobs.

An example of an initial step toward this objective is provided in Figure 4. Job-related postures were assessed by detailed videotape job analyses. Extreme and awkward postures were shown as important risk dimensions that dramatically increase risks for low back disease. In the larger group, which was studied for flexion extremes, these extremes were related to dose (as measured by cumulative time in a given posture).[46] A second example is shown in Figure 5, which presents results from a similar job analysis.[47] This analysis provided evidence that high force and high repetition are associated with increased prevalence of carpal tunnel syndrome. These findings provide hope that common job features can be identified that may prove to be important risk factors.

Objective measurements of musculoskeletal diseases also need to be greatly improved upon. Currently, careful application of precise diagnostic criteria is known to identify only the tip of the iceberg. This limitation results from the poor understanding of the pathophysiology of such conditions. Improvements will require epidemiologists to work with physiologists, kinesiologists, biomechanical engineers, and others in an effort to seek new measures of early effects and to develop test measures or validated symptom complexes that greatly increase the relevant measures of outcome.

In addition to better quantification of both exposure and outcome, more prospective studies of exposure-response relationships are required to guide intervention. For example, critical information needed for job redesign is the type and amount of change in a given risk factor required to predictably achieve a measurable or sufficient reduction in risk of disease. A vexing problem in exposure-response studies relates to the frequency and relatively short latency of many of the work-related musculoskeletal conditions. These features can result in a significant short-term turnover in the work force, which results in significant healthy worker selection effects that must be taken into account.[48]

A final issue deserves comment before I leave the subject of etiologic epidemiology. Despite all the caveats regarding proper assessment of exposure, proper use of exposure data to estimate dose, proper measures of disease, and proper attention to confounding, the results of these studies have had a profound influence on directing social activity on environmental issues. In reaction, there is also an increasing demand that the "perfect" epidemiology study be conducted before action is recommended. It is not for the epidemiologist alone to correct possibly distorted uses of and demands on epidemiology. But these developments place an important burden on each epidemiologist to take a role in proposing responsible public health action to be taken even in the face of scientific uncertainty. Epidemiologists *are* the best-equipped scientists to address the void between the knowledge that exists, the knowledge that is needed, and the public health context in which decisions must be made.[49-51]

Moreover, Sarraci notes that epidemiologic research needs to interact with the disciplines of psychology and sociol-

ogy to develop new and effective ways of communicating risk information to the different component parts of society. Furthermore, he comments, "there is a great need for the research epidemiologist to be more involved in talking and writing to people directly and indirectly through the media. In this way he would overcome his lack of training (shared with other scientists) in putting ideas over in a simple way and would contribute to the development of a soundly informed public."[52] A small step in this direction has recently been taken by in the United States by the National Institute for Occupational Safety and Health (NIOSH), which has developed a formal policy for individual worker notification concerning risks identified in NIOSH's own epidemiology studies.

Intervention Epidemiology

Studies of intervention are important as a means to evaluate hypotheses upon which intervention is based. They also should contribute to evaluating the components and determining the success of intervention programs per se, thus providing support necessary to extend or alter prevention efforts.

An early attempt to identify the impact of intervention in a work environment is provided in Doll's 1958 study, which identified a striking excess of lung and nasal cancer among nickel workers.[53] Believing that a change in industrial process by 1925 had removed the carcinogen, Doll examined time and age patterns for the cancers of interest. He noted some reduction in the excess risk with time, as well as an increase in average age at death, suggesting that there had been higher risk in the past. Unfortunately, lung and nasal sinus cancer have such a long latency that even the 1958 study could not adequately examine the effect of a change that had occurred more than 30 years before; therefore, Doll revisited the question in a follow-up study published in 1970.[54] This time (Table 7) Doll and his colleagues were able to demonstrate convincingly the essential elimination of excess risk for those first employed after the process change in 1925.

In this case, a process change fortuitously eliminated carcinogen exposure and Doll had the wisdom to examine the impact of the removal. A study from Finland on carbon disulfide and cardiovascular disease risk provides an elegant combination of etiologic epidemiology leading to intervention followed by an intervention epidemiology study. The initial evi-

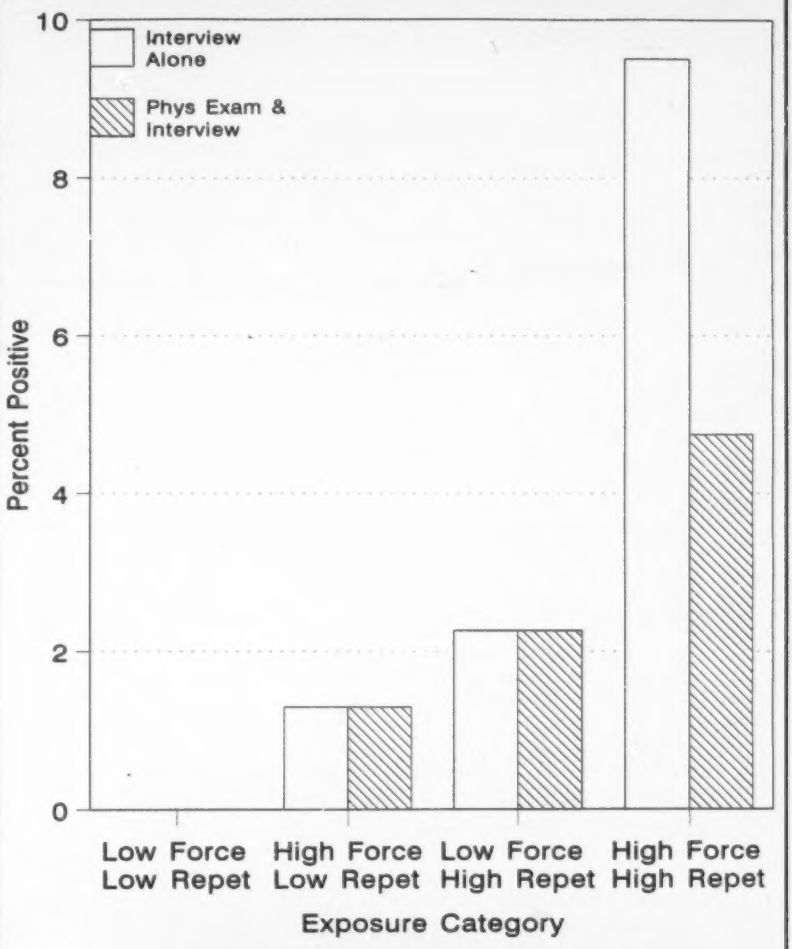


FIGURE 5—Percentage of male subjects with carpal tunnel syndrome (CTS) by force and repetition (Repet) exposure groupings. Data are presented according to whether CTS was diagnosed by interview alone or interview plus a positive physical exam. Adapted with permission from Silverstein et al.,[47] Copyright © 1987, Wiley-Liss.

TABLE 7—Nasal and Lung Cancer Risks among Nickel-Exposed Workers before and after Change in Industrial Process

	First Employed	
	Before 1925	1925–1944
No. workers	563	282
Person-years	5854	4081
Nasal sinus cancer		
Observed	39	0
Expected	0.107	0.036
SMR	364	-
Lung cancer		
Observed	105	8
Expected	13.9	6.1
SMR	7.5	1.3

Note. SMR = standardized mortality ratio. Table adapted from Doll et al.[54]

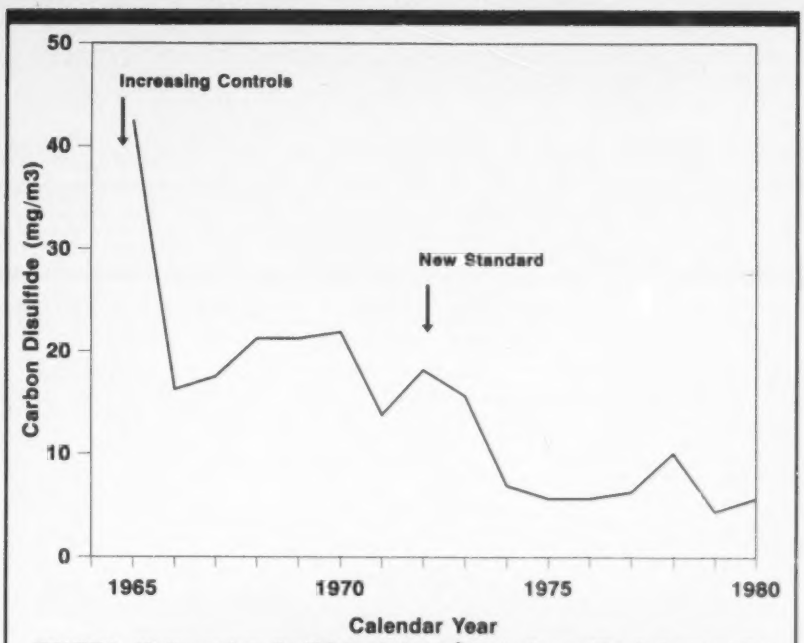


FIGURE 6—Median carbon disulfide levels (mg/m³) in air of rayon fiber factory over the 15-year interval 1965 to 1980. Control efforts were initiated prior to 1965. The exposure limit was reduced from 60 to 30 mg/m³ in 1972. Adapted with permission from Nurminen and Hernberg,[57] Copyright © 1985, *British Journal of Industrial Medicine*.

dence for the risk from carbon disulfide exposure was provided in a cross-sectional study published in 1970.[55] Because exposure control measures had been gradually improving over the years up to 1970, no new action was taken until a 5-year follow-up study demonstrated persistence of the excess risk of death from cardiovascular disease.[56] New and intensive control efforts were initiated and 10 years later a further follow-up study was undertaken. The results show dramatic evidence of risk reduction.[57] The close parallel between exposure reduction and risk reduction (Figure 6) indicates that carbon disulfide risk is reversible. The authors suggest, in fact, that the fraction of prevented or postponed cardiovascular deaths among the exposed was on the order of 70%.

Another intervention study illustrates the positive impact of control measures, although it was not designed to examine for evidence of residual risk. Excess rates of silicosis and tuberculosis were found in Vermont, where granite stonecutting is a major trade. Detailed descriptive studies in the 1930s provided dramatic evidence of work-related disease and resulted in a statewide program to assure appropriate engineering controls at all granite-cutting locations. Twenty years after the control program was started only one new (and questionable) case of silicosis had been documented.[58]

Although this was excellent evidence of the positive impact of the intervention, the study may have asked the wrong question. Fifteen years later a somewhat different question was addressed. Whereas there was no disagreement that the original goal of eliminating *cases of silicosis* had been met, a new epidemiologic study asked whether the intervention had been adequate to eliminate *all pulmonary effects* due to silica. The study examined pulmonary function in currently employed granite workers and revealed that although it was much reduced, exposure-related risk of pulmonary impairment remained. This information is of direct value in setting future intervention goals.[59]

The lung cancer and silicosis examples emphasize again the problem related to prevention efforts when the occupational disease of interest is one with a long latency. For intervention epidemiology as well as the applications discussed earlier, an effective hazard surveillance system could prove particularly valuable.

Recently hazard surveillance was used in the United States to examine the impact of a concerted effort to control silica exposures in silica flour plants. In 1980 a NIOSH investigation found that 23 of 61 employees in two plants had silica-related disease.[60] Using exposure data collected by OSHA, researchers then documented that serious overexposures were present throughout the industry.[61] This second report found that in 89% of the plants surveyed, 25% or more of the dust samples exceeded the standard. Given these findings, NIOSH issued a detailed report on engineering controls.[62] Six years later a follow-up survey of the same plants was carried out to determine the degree to which exposures had been reduced in response to the report. Unfortunately, the results indicated very limited progress that led to new initiatives.[63] This use of hazard surveillance data, however, provided timely evidence that there was a need to carry out more effective primary prevention activity without having to wait until there was evidence from a study of disease rates.

Epidemiologic studies evaluating the impact of occupational controls in industrializing nations should be particularly valuable. Scarce resources are vigorously competed for; researchers must pay special attention to documenting and understanding the use of control technology. It is not surprising, then, that applications of intervention epidemiology are available from these nations as well. For example, Lou and Zhou provided evidence of a more than 70% reduction in silicosis resulting from dust controls introduced to protect a large group of Chinese workers who were exposed to silica.[64] Another illustration of the potential value of such epidemiologic studies comes from a report on a Chinese viscose rayon plant. Evidence is provided that shows increasing production accompanied by increasing exposure levels until the introduction of a solvent recovery system resulted in the achievement of dramatic reductions in exposure.[65]

Finally, there is at least one instance in which a study was designed to examine the effects that were due to several different steps taken to control an occupational disease. The study, by Jarvholm and his colleagues, was initiated in response to the discovery of eight cases of scrotal cancer among lathe operators in a machining plant. The study was designed to review the effects of introducing a series of controls over several years.[66] Because scrotal cancer was a rare event even in this plant, oil acne cases were used as a surrogate disease outcome that was related to the same machining fluid exposures but that occurred with much higher frequency and shorter latency. The effectiveness of the

different control steps could be evaluated by tracking oil acne cases, which had been recorded since 1958. Figure 7 shows that automation (a), free clean underwear (c), and substitution with solvent-refined oils (e) appear to have been the most important control steps. The other controls show limited if any impact.

Epidemiologic studies of intervention efforts are, in principle, no more complicated than those designed for etiology. The lack of a large body of experience to consult makes the task more difficult, but not because the tools are not available. One common impediment appears to be that those responsible for interventions lack the commitment or interest to evaluate whether the intervention is successful. In trying to understand this situation I found Freyman's[67] comments on prevention instructive:

> . . . the ideal result of preventive measures is, in a clinical sense, *zero:* the complete absence of disease. To an activist by nature and education, prevention is dull. It lacks the excitement and satisfaction of discovering pathology and working for a cure. This is reinforced by our custom (in the U.S.) of paying only for performance of a measurable act and placing no dollar value on preventive services rendered to individuals.

In contrast to this characterization of prevention, apt though it may be, I hope I have shown not only that the impact of prevention activities can be exciting, but that it is the missing ingredient in the text of those who would preach prevention and expect to create converts.

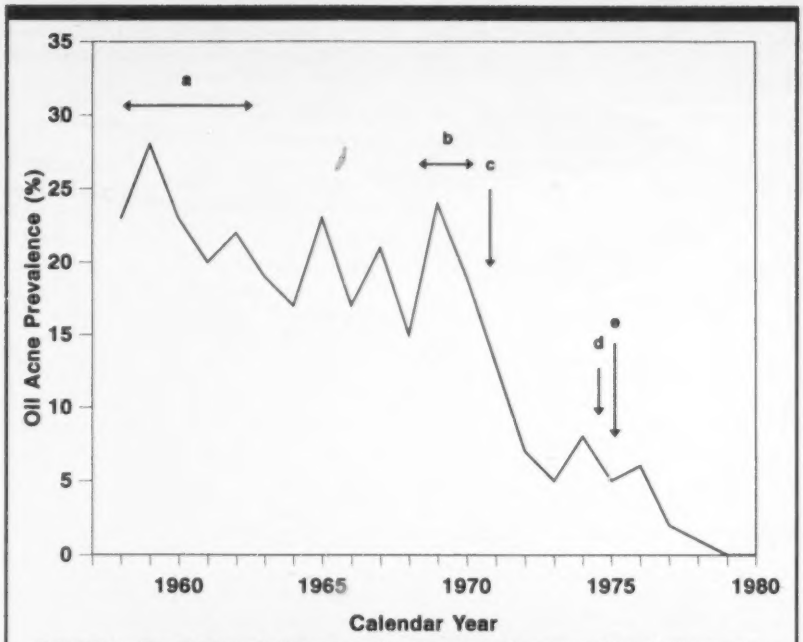


FIGURE 7—Prevalence of oil acne on the forearms of men working in the turning department. Letters refer to major work environment changes: (a) automation of parts removal from machines; (b) further automation allows longer periods away from machines; (c) free underwear provided; (d) free trousers and cotton T-shirts; (e) introduction of solvent-refined instead of acid-refined mineral oils. Adapted with permission from Jarvholm et al.,[66] Copyright © 1982, International Labour Organisation, Geneva.

Conclusion

I have tried to focus this commentary on the full range of opportunities for occupational epidemiology to contribute to prevention. I hope that this paper provokes some further consideration of these opportunities. In closing, however, I feel it is important to reemphasize my concern with the potentially widening gap between epidemiology and public health action. The value of epidemiology to the practice of public health, and in particular the practice of occupational health, is critical. Unfortunately, the academic discipline of epidemiology appears to have become increasingly divorced from applications to prevention in the workplace. The price to be paid is not only financial; the price is also the acceptance of preventable damage to health and loss of life. It is essential that the science of epidemiology and the science of public health be rejoined so that maximum advantage is taken of every opportunity to prevent disease and promote health. □

Acknowledgments

An earlier version of this paper was presented as an invited lecture at the Seventh International Symposium on Occupational Epidemiology held in Tokyo, Japan, in September 1989.

The author would like to express his deep appreciation for the many critical comments and helpful suggestions offered by Ellen A. Eisen, Lawrence J. Fine, Christer Hogstedt, and Linda Rosenstock in the development of the original address and in the final preparation of this document.

References

1. Szklo M. The epidemiologic basis of prevention: how much data do we need? *Johns Hopkins Med J.* 1981;149:64–70.
2. Hernberg, S. Epidemiology in occupational health. In Zenz C, ed., *Occupational Medicine* Yearbook Medical Publishers; 1975.
3. Farr W. Mortality of miners. In Susser M, Adelstein A, eds., *Vital Statistics: A Memorial Volume of Selections from the Reports and Writings of William Farr* Metuchen, NJ: Scarecrow Press; 1975.
4. Dubrow R, Wegman DH. Setting priorities for occupational cancer research and control: synthesis of the results of occupational surveillance studies. *J Natl Cancer Inst.* 1983;71:1123–1142.
5. Kreiss K, Wegman DH, Niles CA, Siroky MB, Krane RJ, Feldman RG. Neurologic dysfunction of the bladder in workers exposed to dimethylaminopropionitrile. *JAMA.* 1980;243:741–745.
6. Froines JR. Inhibition of acetylcholinesterase by neurotoxic aminonitriles. *Biochem Pharmacol.* 1985;34:3185–3187.
7. Froines JR, Postlethwait EM, LaFuente EJ, Liu WCV. In vivo and in vitro release of cyanide from neurotoxic aminopropionitriles. *J Toxicol Environ Health.* 1985;16:449–460.
8. Froines JR, Watson AJ. Evaluation of the inhibition of glycolytic enzymes by the neurotoxin dimethylaminopropionitrile. *J Toxicol Environ Health.* 1985;16:469–479.
9. Aksoy M, Erdem S, DinCol G. Leukemia in shoe-workers exposed chronically to benzene. *Blood.* 1974;44:837–841.
10. Occupational exposure to benzene: final rule. U.S. Congress *Federal Register.* September 11, 1987;52:24460–34578.
11. El Batawi MA, Husbumrer C. Epidemiological approach to planning and development of occupational health services at a national level. *Int J Epidemiol.* 1987;16:288–292.
12. US Dept of Labor, Bureau of Labor Statistics. Annual survey of occupational injuries and illness in the United States, by industry [Data from years 1972–1987]. Washington, DC: US Dept of Labor. Published annually.
13. Baker EL, ed. Surveillance in occupational safety and health. *Am J Public Health.* December 1989; 79(suppl).

14. MacMahon B, Pugh, TF. *Epidemiology.* Boston, Mass: Little Brown & Co Inc; 1970.
15. Froines JR, Wegman DH, Eisen EA. Hazard surveillance in occupational disease. *Am J Public Health.* 1989;79(suppl):26–31.
16. Froines JR, Baron S, Wegman DH, O'Rourke S. Characterization of the airborne concentrations of lead in U.S. industry. *Am J Ind Med.* 1990;18:1–17.
17. Doll R. Epidemiological discovery of occupational cancers. *Ann Acad Med Singapore.* 1984;13(suppl):331–339.
18. Doll R. Mortality from lung cancer in asbestos workers. *Br J Ind Med.* 1955;12:81–86.
19. Breslow L. Industrial aspects of bronchogenic neoplasms. *Dis Chest.* 1955;28: 421–430.
20. McDonald JC, McDonald AD. Epidemiology of asbestos-related lung cancer. In Antman K, Aisner J, eds., *Asbestos Related Malignancy.* Orlando, Fla: Grune & Stratton; 1987:57–79.
21. Dement JM, Harris RL, Symons MJ, Shy CM. Exposures and mortality among chrysotile asbestos workers. Part II. Mortality. *Am J Ind Med.* 1983;4:421–433.
22. McDonald JC, Lidell FDK, Gibbs GW, et al. Dust exposure and mortality in chrysotile mining, 1910–1975. *Br J Ind Med.* 1980;37:11–24.
23. McDonald JC. Asbestos fibers and cancer. *Ann Acad Med Singapore.* 1984;13(suppl): 345–352.
24. Newhouse ML. A study of the mortality of workers in an asbestos factory. *Br J Ind Med.* 1969;26:294–301.
25. Weill H, Hughes J, Waggenspack C. Influence of dose and fiber type on respiratory malignancy risk in asbestos cement manufacturing. *Am Rev Respir Dis.* 1979;120: 345–354.
26. Hammond EC, Selikoff IJ, Seidman H. Asbestos exposure, cigarette smoking habits, and death rates. *Ann N Y Acad Sci.* 1979; 330:473–490.
27. Fuchs S, Valade P. Etudes cliniques et experimentales sur quelques cas d'intoxication par le desmodur T (diisocyanate de toluylène) 1-2-4 et 1-2-6. *Arch Mal Prof.* 1951;12:191–196.
28. Peters JM, Murphy RLH, Pagnotto LD, Van Ganse WF. Acute respiratory effects in workers exposed to low levels of toluene diisocyanate (TDI). *Arch Environ Health.* 1968;16:642–647.
29. Peters JM. Cumulative pulmonary effects in workers exposed to tolylene diisocyanate. *Proc R Soc Med.* 1970;63:14–17.
30. Peters JM, Wegman DH. Epidemiology of toluene diisocyanate (TDI) induced respiratory disease. *Environ Health Perspect.* 1975;11:97–100.
31. Wegman DH, Pagnotto LD, Fine LJ, Peters JM. A dose-response relationship in TDI workers. *J Occup Med.* 1974;16:258–260.
32. Wegman DH, Musk AW, Main DM, Pagnotto LD. Accelerated loss of FEV-1 in polyurethane production workers: a four-year prospective study. *Am J Ind Med.* 1982;3:209–215.
33. Diem JE, Jones RN, Hendrick DJ, et al. Five-year longitudinal study of workers employed in a new toluene diisocyanate manufacturing plant. *Am Rev Respir Dis.* 1982;126:420–428.
34. Weill H, Butcher B, Dharmarajan V, et al. *Respiratory and immunologic evaluation of isocyanate exposure in a new manufacturing plant.* Cincinnati, Ohio: National Institute for Occupational Safety and Health; 1981. US DHHS publication NIOSH 81-125.
35. Kristensen TS. Cardiovascular diseases and the work environment. *Scand J Work Environ Health.* 1989;15:165–179.
36. Karasek RA, Theorell T, Schwartz JE, Schnall PL, Pieper CF, Michela JL. Job characteristics in relation to the prevalence of myocardial infarction in the US Health Examination Survey (HES) and the Health and Nutrition Examination Survey (HANES). *Am J Public Health.* 1988;78: 910–918.
37. Kristensen TS. Cardiovascular diseases and the work environment. *Scand J Work Environ Health.* 1989;15:165–179.
38. LaCroix A, Haynes S. Occupational exposure to high demand/low control work and coronary heart disease incidence in the Framingham cohort. *Am J Epidemiol.* 1984;120:481.
39. Siegrist J, Matschinger H, Cremer P, Seidel D. Atherogenic risk in men suffering from occupational stress. *Atherosclerosis.* 1988;69:211–218.
40. Knutsson A, Åkerstedt T, Jonsson BG, Orth-Gomer K. Increased risk of ischaemic heart disease in shift workers. *Lancet.* 1986;2:89–92.
41. Paffenbarger RS, Hale WE, Brand RJ, Hyde RT. Work-energy level, personal characteristics, and fatal heart attack: a birth cohort effect. *Am J Epidemiol.* 1977; 105:200–213.
42. Nygard C-H, Luopajarvi T, Ilmarinen J. Musculoskeletal capacity of middle-aged women and men in physical, mental and mixed occupations: a 3.5 year follow-up. *Eur J Appl Physiol.* 1988;57:181–188.
43. Ilmarinen J. Physiological criteria for retirement age. *Scand J Work Environ Health.* 1988;14(suppl):1,88–89.
44. Silverstein BA, Fine LJ, Armstrong TJ. Carpal tunnel syndrome: causes and a preventive strategy. *Semin Occup Med.* 1986; 1:213–221.
45. Hagberg M, Wegman DH. Prevalence rates and odds ratios of shoulder-neck diseases in different occupational groups. *Br J Ind Med.* 1987;44:602–610.
46. Punnett L, Fine LJ, Keyserline WM, Herrin GD, Chaffin DB. Back disorders and nonneutral trunk postures of automobile assembly workers. *Scand J Work Environ Health* 1991;17:337–346.
47. Silverstein BA, Fine LJ, Armstrong TJ. Occupational factors and carpal tunnel syndrome. *Am J Ind Med.* 1987;11:343–358.
48. Punnett L, Robins JM. Adjusting for healthy worker selection effect in cross-sectional studies. *Scand J Work Environ Health.* 1987;13:169.
49. Smith AH. Epidemiologic input to environmental risk assessment. *Arch Environ Health.* 1988;43:124–127.
50. Whittemore AS. Epidemiology in risk assessment for regulatory policy. *J Chronic Dis.* 1986;39:1157–1168.
51. McMichael AJ. Setting environmental exposure standards: the role of the epidemiologist. *Int J Epidemiol.* 1989;18:10–15.
52. Saracci R. Epidemiological strategies and environmental factors. *Int J Epidemiol.* 1978;7:101–111.
53. Doll R. Cancer of the lung and nose in nickel workers. *Br J Ind Med.* 1958;15:217–223.
54. Doll R, Morgan LG, Speizer FE. Cancer of the lung and nasal sinuses in nickel workers. *Br J Cancer.* 1970;24:623–632.
55. Hernberg S, Partanen T, Nordman C-H, Sumari P. Coronary heart disease among workers exposed to carbon disulfide. *Br J Ind Med.* 1970;27:313–325.
56. Tolonen M, Hernberg S, Nurminen M, Titola K. A follow-up study of coronary heart disease in viscose rayon workers exposed to carbon disulfide. *Br J Ind Med.* 1975;32: 1–10.
57. Nurminen M, Hernberg S. Effects of intervention on the cardiovascular mortality of workers exposed to carbon disulphide: 15 year follow up. *Br J Ind Med.* 1985;42:32–35.
58. Hosey AD, Trasko VM, Ashe HB. *Control of silicosis in Vermont granite industry—progress report.* Washington, DC: US Public Health Service, Bureau of State Services; 1957. US DHEW publication 557.
59. Theriault GP, Peters JM, Johnson WM. Pulmonary function and roentgenographic changes in granite dust exposure. *Arch Environ Health.* 1974;28:23–27.
60. Centers for Disease Control. Silicosis—Illinois. *MMWR.* 1980;29:205–206.
61. Banks DE, Morring KL, Boehlecke BA. Silicosis in the 1980's. *Am Ind Hyg Assoc J.* 1981;42:77–79.
62. *Health hazard control technology assessment of the silica flour milling industry.* Cincinnati, Ohio: US Public Health Service; 1984. US DHHS publication NIOSH 184–110.
63. Centers for Disease Control. Exposure trends in silica flour plants—United States, 1975–86. 1989;38:380–383.
64. Lou J, Zhou C. Prevention of silicosis and prediction of its future prevalence in China. *Am J Public Health.* 1989;79:1613–1616.
65. Liang Y-X, Qu D-Z. Cost-benefit analysis of the recovery of carbon disulfide in the manufacturing of viscose rayon. *Scand J Work Environ Health.* 1985;11(suppl): 4,60–63.
66. Jarvholm B, Fast K, Lavenius B, Tomsic P. Prevention of scrotal cancer induced by mineral oils in an engineering industry. In *Prevention of Occupational Cancer—International Symposium,* Occupational and Health Series 46. Geneva: International Labour Organisation; 1983.
67. Freyman JG. The origins of disease orientation in American medical education. *Prev Med.* 1981;10:663–673.

Chlorination, Chlorination By-products, and Cancer: A Meta-analysis

ABSTRACT

Objectives. Individual epidemiological investigations into the association between chlorination by-products in drinking water and cancer have been suggestive but inconclusive. Enough studies exist to provide the basis for a meaningful meta-analysis.

Methods. An extensive literature search was performed to identify pertinent case-control studies and cohort studies. Consumption of chlorinated water, surface water, or water with high levels of chloroform was used as a surrogate for exposure to chlorination by-products. Relative risk estimates were abstracted from the individual studies and pooled.

Results. A simple meta-analysis of all cancer sites yielded a relative risk estimate for exposure to chlorination by-products of 1.15 (95% CI: 1.09, 1.20). Pooled relative risk estimates for organ-specific neoplasms were 1.21 (95% CI: 1.09, 1.34) for bladder cancer and 1.38 (95% CI: 1.01, 1.87) for rectal cancer. When studies that adjusted for potential confounders were pooled separately, estimates of relative risks did not change substantially.

Conclusions. The results of this meta-analysis suggest a positive association between consumption of chlorination by-products in drinking water and bladder and rectal cancer in humans. (*Am J Public Health.* 1992;82:955–963)

Robert D. Morris, MD, PhD, Anne-Marie Audet, MD, MS, Italo F. Angelillo, DDS, MPH, Thomas C. Chalmers, MD, and Frederick Mosteller, PhD

Introduction

After its first use in the Chicago stockyards in 1908, chlorination of drinking water spread rapidly throughout the United States and produced dramatic reductions in morbidity and mortality associated with waterborne disease.[1] Since then, chlorine has remained the method of choice for water purification; it is currently added to approximately 75% of the nation's drinking water.[2] In 1974, Rook[3] discovered that the combination of chlorine with organic compounds in drinking water produced halogenated organic compounds—chloroform, in particular. In 1975, the National Organics Reconnaissance Survey[4] verified that chlorination by-products were a major contaminant of chlorinated surface water. The study also showed that chloroform concentration correlates highly with the concentrations of other halogenated hydrocarbons and is, therefore, a useful indicator for the presence of these compounds. Since this finding, a plethora of epidemiological studies[5–27] have investigated possible associations between consumption of chlorinated water and cancer. These studies consider a wide range of populations and regions and demonstrate somewhat inconsistent patterns of association. We used meta-analytic methods to pool the results of these studies in an attempt to further our understanding of the relationship between water chlorination and neoplastic diseases. Meta-analysis provides greater statistical power and greater resolution in the estimation of relative risks than do individual studies.

Methods

Identification of Relevant Studies

To identify references pertaining to the relationship between chlorination of drinking water and neoplastic diseases, we used the Medline data retrieval system to search the medical, public health, and biological literature from 1966 through 1991. Relevant papers were obtained and reviewed to locate additional references. Because initial analyses suggested an association between chlorination and both bladder and rectal cancer, the Medline system was used to conduct a second search of the literature over the same 25-year period for papers that considered these neoplasms and mentioned water in the title, abstract, or keywords.

Only those studies that identified morbidity or mortality as well as exposure and potential confounders at the level of the individual (i.e., case-control or cohort studies) were included in the meta-analysis. Studies that considered incidence and exposure at the level of a region or community (i.e., ecological studies) were excluded.

Quality Scoring and Exposure Assessment

Each article was blinded as to authors, institutions, journal, and study re-

Robert D. Morris is with the Division of Biostatistics, Medical College of Wisconsin, Milwaukee, Wis. Anne-Marie Audet is with the American College of Physicians, Philadelphia, Pa. Italo F. Angelillo is with the Institute of Hygiene and Preventive Medicine at the Universita Degli Studi di Napoli Fredrico II, Naples, Italy. Thomas Chalmers and Frederick Mosteller are with the Technology Assessment Group, Harvard University School of Public Health, Boston, Mass.

Requests for reprints should be sent to Robert D. Morris, MD, PhD, Division of Biostatistics, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226.

This paper was submitted to the Journal August 21, 1991, and accepted with revisions March 23, 1992.

sults. Two independent readers scored each paper for quality, using a scoring system that incorporates elements of systems developed by Chalmers et al.[28] and Longnecker et al.[29] Studies were scored on the basis of selection of subjects, measurement of and adjustment for confounding variables, exposure assessment, and statistical analysis. The overall quality score was calculated from three subscores: a general methods score, a data analysis score, and an exposure assessment score. Each subscore was calculated as the percentage of applicable quality criteria that were met in each study. The criteria employed are listed with the results. The cumulative quality score was a weighted average of the three scores, with both general methods and exposure assessment receiving twice the weight of the data analysis score.

Exposure assessment is the most difficult task in a study of this kind. Three factors determine a person's cumulative exposure to an environmental contaminant or risk factor: (1) the environment to which the person was exposed, (2) the level of the proposed agent present in the environment, and (3) the degree to which the person is exposed to that environment. In the current context, these exposures refer to (1) the source of tap water, (2) concentrations of chlorination by-products in that tap water, and (3) the amount of tap water consumed. Although exposure at a fixed point in time (e.g., the time of diagnosis or death) often correlates highly with lifetime exposure, a complete assessment of exposure must also include the historical records for each of these three factors. Thus there are six factors that must be evaluated in each study: three exposure factors plus historical records for each of these factors. The quality of exposure assessment was scored for each study as the percentage of these factors that were evaluated in the study.

Most of the identified studies used consumption of surface water as an indirect measure of exposure to chlorination by-products. Unchlorinated natural waters do not contain significant amounts of chlorinated hydrocarbons. Consequently, the comparison of consumers of chlorinated water with consumers of unchlorinated water provided an acceptable surrogate for comparing exposure vs nonexposure to chlorination by-products.

Similarly, chlorinated surface water contains more chlorination by-products than does groundwater. Whereas virtually all surface water is chlorinated, groundwater tends to be less heavily chlorinated, and in the case of private wells the water is often unchlorinated. In addition, groundwater contains less organic matter than does surface water. Organic matter combines with chlorine to form chlorination by-products. Measurements of chlorination by-products in the drinking water supply[3] have shown that chlorinated surface water contains much higher levels of these by-products than does chlorinated groundwater (medians of 50.7 and 0.8 ppb, respectively) even when the surface water is drawn from protected reservoirs. For these reasons, consumption of surface water vs groundwater was also used as a surrogate for exposure to chlorination by-products.

Although most of the studies used either consumers of groundwater or consumers of unchlorinated water for the control group, one study[22] used consumers of chloraminated water as the control group. Chloramination is a process whereby both chlorine and ammonia are added to filtered drinking water to form chloramines. This chloramine residual is intended to provide sufficient bactericidal capacity to maintain water purity in the distribution system without introducing large amounts of free chlorine. This is a process that minimizes the formation of chlorination by-products. Therefore, this comparison of chlorinated with chloraminated water consumption provided a surrogate for exposure vs nonexposure to chlorination by-products.

Chloroform is the most common chlorination by-product and consequently has been the focus of health concerns related to water chlorination. Several studies included measured chloroform levels in their exposure assessments. Although recent research[30] suggests that other by-products may be more toxic, chloroform concentration correlates well with the concentration of other by-products.[4] Therefore, measures of chloroform concentrations were deemed to be the most accurate measures of chlorination by-products.

Extraction of Relative Risk Estimates

The odds ratios or relative risks for cancer among consumers of drinking water containing chlorination by-products were identified for each of the selected studies. Exposure was defined on the basis of the categorization of subjects as consumers of chlorinated water or surface water in the individual studies; nonexposure was defined on the basis of the categorization of subjects as consumers of nonchlorinated water, chloraminated water, or groundwater. When risks were stratified by sex, they were recorded as such. The variance of the natural log of the odds ratio or relative risk was derived with the statistics provided in the particular study. If an odds ratio was reported to be different from unity with P less than a specified significance level, the variance was calculated assuming P equal to this maximum value (e.g., given $P < .05$, $P = .05$ was assumed). Calculation of a variance on the basis of this conservative assumption resulted in the maximum possible variance and minimized the probability of a type I error in the meta-analyses. If a study reported a coefficient from a logistic regression, the regression coefficient was used along with an estimate of average exposure for the exposed sample to calculate an estimate for the relative risk. Studies that stratified according to level of exposure were analyzed by pooling the values from each stratum to yield a single value from each study estimating the relative risk for the exposed vs nonexposed populations.

The neoplasms considered in these studies were sufficiently rare that odds ratios were deemed to be close approximations of relative risks. Therefore, we did not distinguish between odds ratios and relative risks in our meta-analysis. All results listed may be considered as relative risk estimates.

Some studies reported morbidity data and others reported mortality data. For relative risks for mortality to differ from relative risks for morbidity, the exposure in question must alter disease survival. We have no reason to expect that water chlorination alters disease survival. To evaluate this possible bias, we analyzed the two groups separately as part of the sensitivity analysis described below.

Meta-analysis

The relative risk estimates were analyzed in two stages, a meta-analysis of all studies and a meta-analysis of selected subgroups of studies. In the first stage, meta-analyses were performed for each cancer site by pooling all pertinent studies without stratifying for sex. If a given tumor site was evaluated in at least two studies, relative risk estimates were pooled and analyzed for that site. The random effects model described by DerSimonian and Laird[31] was used to combine the collected values. This procedure yielded a single estimate of the relative risk for or-

gan-specific neoplasms in individuals exposed to chlorination by-products in drinking water.

Meta-analysis, in addition to improving our ability to identify small but significant relative risks in the context of apparently contradictory research findings, enhances our capacity to interpret truly negative results. Statistical power is a measure of the probability of finding a significant difference, assuming a specified effect and a specified type I error rate. A detailed analysis of statistical power was carried out for each cancer site. For each negative result ($P > .05$) the probability of detection of relative risks of 1.2, 1.4, and 1.6 was determined on the basis of the calculated relative risk estimate and variance.

Sensitivity and Dose-Response Analyses

If the first analysis demonstrated a significant association ($P < .05$) between chloroform exposure and neoplastic disease for a given site, a second set of more detailed analyses was performed on relative risk estimates pertaining to that site. The influence of potential confounding variables, sex differences, quality score, and accuracy of exposure assessment on the effect of chlorination by-products were all evaluated in this stage. In addition, studies that provided dose-response relationships were pooled separately.

The effect of inaccurate determination of exposure to chlorination by-products (misclassification bias) and the effect of incomplete control for confounding variables were of particular concern. The authors of each study chose to adjust for a slightly different set of potential confounding variables. A failure to adjust for some factor (e.g., smoking in bladder cancer) might well lead to erroneous estimates of risk. To determine whether such differences among studies might have an impact on the pooled results, individual studies that controlled for confounders were grouped for separate meta-analyses. The relative risk estimates of cancer for each site were also pooled separately for each sex. Studies were also divided into two groups and analyzed according to their score for overall quality and for quality of exposure assessment. In each case, studies with scores above the mean were compared with studies whose scores were below the mean.

A conventional literature review pools findings implicitly by weighting according to a subjective and unspecified quality scoring system. Although weighting by quality score is intuitively appealing, there is no established protocol either for calculating a quality score or for weighting according to score. As a simple test, we repeated the meta-analyses for bladder and rectal cancer, using quality scores as weights.

Meta-analysis was also used to evaluate dose-response relationships. For each study that stratified by level of exposure, relative risk estimates for high, medium, and low levels of exposure were determined. The studies were then combined, with each level pooled separately.

The dose-response analysis excluded those studies that evaluated exposure as a dichotomous variable. In order to determine the impact of this exclusion, we repeated the dose-response analysis with these studies included. In order to include these studies, we assumed that the exposed group in studies with dichotomous exposure measures had experienced moderate levels of exposure.

Results

Literature Search

The literature search identified 10 case-control studies[15–24] and two cohort studies[25,26] that investigated the relationship between some measure of exposure to chlorination by-products in the drinking water and cancer. One case-control study[19] considered source of drinking water as one of many potential risk factors for bladder cancer. The association between bladder cancer and drinking water was stated to be nonsignificant, but numerical results were not reported, making the inclusion of this study in the meta-analysis impossible. One cohort study considered bladder cancer in a group of workers with exposure to a known carcinogen. This cohort was deemed not to represent the general population and was also excluded. The 10 studies that were entered into the meta-analysis are listed in Table 1.

One study[22] evaluated a variety of neoplasms and found an association between bladder cancer and chlorination. It was followed by a second study[27] that analyzed the same sample with more thorough methods for assessment of exposure to chlorination by-products. The second study reported two odds ratios, one comparing high exposure to no exposure and the other comparing moderately high to moderately low exposure. The second of these two odds ratios was used in the analyses of bladder cancer except in the dose-response analysis, in which both values were used.

Quality Scoring

Table 2 shows the results of the quality scoring procedure. All 10 studies identified the source of drinking water at the residences of the subjects either at the time of the study or at the time of death. Chloroform concentrations and specific tap water consumption rates for the subjects were determined in a minority of the studies (4 in both cases). Historical patterns of drinking water consumption were not quantified in any of the studies. Historical sources of tap water were evaluated in 6 of the studies. No study used historical values of chlorination by-product concentrations, because these are not generally available, but 7 used some portion of the historical record, such as water utility chlorination logs, to estimate historical levels of chlorination by-products in drinking water. The studies are listed in Table 1 in order of their quality scores.

Selection of cases and controls followed accepted procedures and was not a major source of bias in these studies. Information bias associated with failure to adjust for confounding variables may have been more of a problem. With the exception of age (100%), sex (100%), population density (80%), and race (60%), potentially confounding variables were not generally adjusted for. Statistical analyses in these studies were usually thorough with regard to the determination of relative risks, but they tended to be less rigorous with regard to analyzing other aspects of the sample groups. Although most studies listed P values (70%), adjusted for confounders (90%), and identified specific tests used (80%), few listed or analyzed demographic data (40%). Only one study (10%) provided power calculations.

Extraction of Risk Estimates and Meta-analysis

Table 3 lists the relative risk estimates extracted from each study. Although 39 of these 64 estimates (64%) are greater than 1.00, only 19 (30%) are significantly greater than 1.00 ($P < .05$).

Table 4 lists the results of the organ-specific meta-analyses, with power calculations. Bladder cancer has an overall relative risk estimate of 1.21 (95% CI: 1.09, 1.34). Rectal cancer has a relative risk estimate of 1.38 (95% CI: 1.01, 1.87). Lung, breast, brain, kidney, liver, pancreatic, esophageal, colonic, and gastric neoplasms all have pooled estimates of relative risk that are not significantly different from unity.

Note that all 12 relative risk estimates were greater than 1.00, suggesting that the

TABLE 1—Summary of Studies Included in the Meta-analysis Relating Chlorination By-products in Drinking Water to Cancer

Authors	Year[a]	Population	Cases	Controls	Exposure Assessment	Covariates	Analysis
Cantor et al.[20]	1987 (78)	White residents of 10 US regions	2982 newly diagnosed cases of bladder cancer	Community controls matched 2:1 for age, sex, and region by random dialing and Medicare files	History of residence, and beverage consumption combined with survey and sampling of water utilities	Age, sex, cigarette consumption, occupation, population of place of longest residence	Logistic regression
Young, Wolf, Kanarek[24]	1987 (75)	Wisconsin residents	347 newly diagnosed cases of colon cancer	General population controls matched 2:1 for age and sex	Water consumption by interview and chloroform by historical records and measurement	Age, sex, population of place of residence	Logistic regression
Zierler, Danley, Feingold[22,27b]	1986, 1988 (22, 71)	Massachusetts residents	51 645 kidney, bladder, stomach, pancreas, colon, lung, breast cancer deaths	Deaths from cardiovascular, cerebrovascular, or pulmonary disease or lymphatic cancer by death certificate	Chlorination/ chloramination at address on death certificate	Age, sex, surface/ groundwater, population density, and level of poverty in county of residence	Odds ratios with Mantel-Haenszel adjustment
Cragle, Shy, Struba, Siff[21]	1985 (65)	Patients of 7 North Carolina hospitals	200 newly diagnosed cases of colon cancer	Noncancer hospital controls matched 2:1 for age, sex, vital status, and hospital	Years of exposure to chlorinated water by residential history and survey of water utilities	Age, sex, genetic risk, dietary fiber, region, urban residency, smoking, alcohol use, education, number of pregnancies	Logistic regression
Lawrence, Taylor, Trock, Reilly[23]	1984 (62)	New York female teachers	395 colorectal cancer deaths	Noncancer deaths matched 1:1 for age and sex by death certificate	20-year employment and residential history with model for chloroform levels	Age, marital status, population density, type of surface water	Logistic regression
Wilkins, Comstock[25c]	1981 (61)	31 000 residents of Washington County, MD	Diagnosis of cancer in 12 years following initial survey	*Unexposed:* Deep well users	*Exposed:* Users of chlorinated surface water in Hagerstown, Md	Age, sex, smoking, marital status, education	Logistic regression
Gottlieb, Carr, Clarkson[18]	1982 (49)	Louisiana residents	10 205 kidney, bladder, stomach, liver, and colorectal cancer deaths	Two groups of cancer deaths and one of noncancer deaths matched 1:1 for age, sex, race, and parish by death certificate	Chlorinated/ unchlorinated water at address on death certificate and at birthplace	Groundwater/surface water, cardiovascular disease	Stratified odds ratio with χ^2 test
Brenniman, Lagos, Amsel, Namekata, Wolf[19]	1980 (46)	White residents of Illinois	3208 gastro-intestinal and urinary tract cancer deaths	Noncancer deaths matched 14:1 for age, sex, and county by death certificate	Chlorinated/ unchlorinated water at address on death certificate	Urbanicity, population density	Odds ratios with Mantel Haenszel adjustment
Young, Kanarek, Taisle[17]	1961 (45)	White female residents of Wisconsin	8029 cancer deaths	Noncancer deaths matched 1:1 for age, sex, race, and county by death certificate	Chlorinated/ unchlorinated water at address on death certificate	Population density, occupational risk, marital status, rural runoff in water	Logistic regression
Alvanja, Goldstein, Susser[15]	1978 (43)	NY State residents	3446 gastro-intestinal and urinary tract cancer deaths	Noncancer and lung cancer deaths matched 1:1 for age, sex and county by death certificate	Chlorinated/ unchlorinated water at address on death certificate	Urbanicity, occupation	Stratified odds ratio with χ^2 test

Note. Studies are listed in descending order with respect to the quality score. (Scores are presented in parentheses after the year.)
[a]Overall quality scores that are listed with the year may range from 0 to 100.
[b]The second Zierler study[27] was conducted on the same population, but was limited to bladder cancer. Note that this study had a substantially higher quality score than the earlier study (71 vs 22).
[c]Wilkins and Comstock's study was a cohort study. All other studies listed were case-control studies.

association of chlorination by-products with cancer, although modest, is more general than the two statistically significant sites. Pooling all 12 sites yields an estimate of the overall relative risk for cancer associated with chlorination by-products in drinking water of 1.15 ($P < .0001$). Eliminating bladder, rectal, and colorectal cancer and repeating this overall meta-analysis results in a relative risk estimate of 1.09 ($P = .006$).

The power of the negative meta-analyses to detect a relative risk of 1.2 at a significance level of .05 ranges from .27 for breast cancer to .70 for pancreatic cancer, compared with a power of only .06 for brain tumors. For a relative risk greater than 1.6, power for all of the sites with negative findings rises above .90, with the exception of that for brain cancer, which only reaches .18.

Sensitivity and Dose-Response Analyses

Of the sites listed, only bladder and rectal cancer met the established criteria for further analysis. In addition, colonic neoplasms were included in these analyses to evaluate the apparent differences between risks for colonic and rectal neoplasms given the anatomic and physiologic relationship between the two sites. For bladder cancer, the effects of adjusting for smoking, occupation, or population density were evaluated. For rectal and colon cancer, only population density was analyzed (other risk factors, such as diet, were not included in the studies selected). Table 5 lists the results of these analyses.

For bladder cancer, adjustments for confounders do not appear to alter the association with exposure to chlorination by-products. Separate meta-analyses of studies that adjusted for population density, smoking, or occupation yielded relative risk estimates of 1.20, 1.25, and 1.22, respectively. The single study that adjusted for all three factors had a relative risk estimate of 1.19. Estimates of relative risk for bladder cancer were 1.17 for females and 1.24 for males.

The bladder cancer studies with quality scores above the average for this group had a pooled relative risk estimate of 1.25 (95% CI: 1.08, 1.45). Those with below-average quality scores had a pooled value of 1.16 (95% CI: 0.96, 1.40). Improved exposure assessment increases the estimate of relative risk and narrows the confidence limits around this estimate from 1.16 (95% CI: 0.96, 1.40) to 1.25 (95% CI 1.09, 1.45).

TABLE 2—Items Used in Quality Scoring for Case-Control and Cohort Studies of the Association between Consumption of Chlorination By-products in Drinking Water and Cancer

Quality Scoring Item	Studies Complying, %[a]
Case and control selection	
Random selection of cases	90
Selection blinded to exposure status	70
Specific disease criteria given	100
Disease validated by histology or other "gold standard"	40
No known association between control status and exposure	70
Random selection of controls	80
Adjustment or matching for confounders	
Age	100
Sex	100
Population density	90
Race	60
Smoking	40
Occupation	40
Diet	10
Exposure assessment	
Drinking water source for study subjects identified	100
Tap water consumption for study subjects assessed	40
Chloroform in drinking water sources measured	40
Historical drinking water sources for subjects evaluated	80
Historical patterns in tap water consumption for subjects evaluated	0
Historical chlorination levels in water sources evaluated	70
Data analysis	
Power calculation performed	10
Demographic data listed	40
Statistical analysis of demographic data	10
Correlations between independent variables listed	20
Precise *P* values listed	70
Test specified	80

[a]If compliance is not specifically indicated in the text, noncompliance is assumed.

For rectal cancer, adjustments for confounders do not appear to alter the association with exposure to chlorination by-products. Separate meta-analysis of studies that adjusted for population density yielded a relative risk estimate of 1.33. Estimates of relative risk for rectal cancer were 1.11 for females and 1.24 for males.

The rectal cancer studies with quality scores above the average for this group had a pooled relative risk estimate of 1.91 (95% CI: 1.56, 2.35). Those with below-average quality scores had a pooled value of 1.07 (95% CI: 0.89, 1.29). Improved exposure assessment increases the estimate of relative risk from 1.23 (95% CI: 0.89, 1.70) to 1.57 (95% CI: 1.09, 2.24).

A different picture emerges with regard to colon cancer. No grouping of studies, either by sex or by adjustment for population density, results in a significant estimate of relative risk for colon cancer. In the quality score breakdown, high-quality-score studies had lower risk estimates than did low-quality-score studies (1.10 [95% CI: 0.79, 1.53] vs 1.13 [95% CI: 0.86, 1.48]).

Weighting by quality score alone produced relative risks of 1.34 for bladder cancer and 1.46 for rectal cancer (95% CIs: [1.20, 1.49] and [1.29, 1.64], respectively).

Estimates of relative risk for bladder, colon, and rectal cancer at low, medium, and high exposure levels are shown in Table 6. The values for bladder cancer rise steadily from 1.03 to 1.20 to 1.41 and demonstrate increasing levels of significance with increasing exposure level. Rectal cancer also shows a pattern of increasing relative risk with increasing cumulative exposure from 1.13 to 1.29 to 2.04. The inclusion of studies with dichotomous measures of exposure produces estimates of relative risk for moderate exposure that fall, in each case, between the estimates for high and low exposure.

A comparison between studies evaluating incidence and those evaluating mortality did not show a substantial difference between the two categories. For bladder cancer, the pooled relative risk estimates were 1.22 (95% CI: 1.03, 1.44) for the five mortality studies and 1.24 (95% CI: 0.92, 1.67) for the two incidence stud-

TABLE 3—Estimates of the Relative Risk for Neoplasia Associated with the Consumption of Chlorination By-products in Drinking Water as Extracted from Individual Studies and Grouped by Cancer Site

Author	Year	Site	Sex	Relative Risk Estimate	95% CI
Alvanja	1978	Bladder	MF	1.69	1.07, 2.67
Brenniman	1980	Bladder	MF	0.98	0.77, 1.25
Young	1981	Bladder	F	1.15	0.70, 1.89
Gottlieb	1982	Bladder	MF	1.18	0.95, 1.45
Wilkins	1986	Bladder	MF	2.20	0.71, 6.82
Cantor	1987	Bladder	MF	1.19	1.07, 1.32
Zierler	1988	Bladder	MF	1.40	1.20, 2.10
Young	1981	Brain	F	2.13	1.13, 4.04
Gottlieb	1982	Brain	MF	0.85	0.67, 1.09
Young	1981	Breast	F	1.25	0.99, 1.58
Gottlieb	1982	Breast	F	1.19	1.01, 1.42
Zierler	1986	Breast	F	0.89	0.85, 0.93
Wilkins	1986	Breast	F	2.27	1.16, 4.44
Alvanja	1976	Colon	MF	1.81	1.15, 2.24
Brenniman	1980	Colon	MF	1.11	0.99, 1.25
Young	1981	Colon	F	1.52	1.25, 1.85
Gottlieb	1982	Colon	MF	1.01	0.84, 1.20
Cragle	1985	Colon	MF	0.10	0.01, 0.79
Zierler	1986	Colon	MF	0.89	0.86, 0.92
Wilkins	1986	Colon	MF	0.89	0.57, 1.39
Alvanja	1978	Colorectal	MF	1.71	1.31, 2.24
Brenniman	1980	Colorectal	MF	1.13	1.02, 1.26
Young	1981	Colorectal	F	1.45	1.22, 1.73
Gottlieb	1982	Colorectal	MF	1.40	0.73, 2.69
Lawrence	1984	Colorectal	F	0.97	0.74, 1.27
Zierler	1986	Colorectal	MF	0.92	0.85, 0.99
Wilkins	1986	Colorectal	MF	1.03	0.67, 1.58
Young	1987	Colorectal	MF	0.90	0.60, 1.35
Alvanja	1978	Esophagus	MF	2.12	1.10, 4.02
Brenniman	1980	Esophagus	MF	0.97	0.68, 1.38
Young	1981	Esophagus	F	0.95	0.33, 2.76
Gottlieb	1982	Esophagus	MF	1.08	0.78, 1.48
Wilkins	1986	Esophagus	MF	1.76	0.53, 5.84
Young	1981	Kidney	F	0.99	0.69, 1.41
Gottlieb	1982	Kidney	MF	1.49	1.12, 1.99
Zierler	1986	Kidney	MF	1.00	0.90, 1.11
Wilkins	1986	Kidney	MF	2.76	0.67, 11.37
Brenniman	1980	Liver	MF	1.00	0.58, 1.72
Young	1981	Liver	F	1.05	0.75, 1.48
Gottlieb	1982	Liver	MF	1.20	0.93, 1.54
Wilkins	1986	Liver	MF	2.98	0.92, 9.65
Alvanja	1978	Lung	MF	1.79	1.06, 2.97
Young	1981	Lung	F	0.80	0.61, 1.04
Gottlieb	1982	Lung	MF	1.12	0.93, 1.34
Zierler	1986	Lung	MF	0.94	0.91, 0.97
Wilkins	1986	Lung	MF	0.96	0.61, 1.51
Alvanja	1978	Pancreas	MF	1.97	1.23, 3.16
Brenniman	1980	Pancreas	MF	1.02	0.85, 1.23
Young	1981	Pancreas	F	1.15	0.81, 1.64
Gottlieb	1982	Pancreas	MF	1.07	0.86, 1.30
Zierler	1986	Pancreas	MF	0.94	0.89, 0.99
Wilkins	1986	Pancreas	MF	0.80	0.44, 1.45
Alvanja	1978	Rectum	MF	1.93	1.22, 3.05
Brenniman	1980	Rectum	MF	1.22	0.95, 1.56
Young	1981	Rectum	MF	1.21	0.82, 1.78
Gottlieb	1982	Rectum	MF	1.96	1.54, 2.50
Zierler	1986	Rectum	MF	0.96	0.89, 1.04
Wilkins	1986	Rectum	MF	1.42	0.70, 2.88

TABLE 4—Relative Risk Estimates and Power Calculations from the Site-Specific Meta-analyses of the Associations between Exposure to Chlorination By-products and Cancer

Site	n[a]	Relative Risk Estimate	95% CI	P	Power[b] for Detection of Specified Relative Risks ($\alpha = .05$) 1.20	1.40	1.60
Bladder	7	1.21	1.09, 1.34	<.0001	...	...	...
Brain	2	1.29	0.53, 3.14	.56	.06	.11	.18
Breast	4	1.18	0.90, 1.54	.24	.27	.69	.93
Colon	7	1.11	0.91, 1.35	.32	.48	.83	>.99
Colorectal	8	1.15	0.97, 1.37	.10	.54	.97	>.99
Esophagus	5	1.11	0.85, 1.45	.43	.27	.69	.93
Kidney	4	1.16	0.89, 1.51	.23	.27	.71	.94
Liver	4	1.15	0.94, 1.40	.16	.44	.92	>.99
Lung	5	1.01	0.86, 1.18	.94	.27	.69	.93
Pancreas	6	1.05	0.91, 1.22	.48	.70	>.99	>.99
Rectum	6	1.38	1.01, 1.87	.04	...	...	...
Stomach	6	1.14	0.94, 1.38	.19	.46	.93	>.99

[a]Number of studies evaluating specific cancer site.
[b]Power is defined as $1 - \beta$, where β is the probability of type II error at the given levels of true relative risk and significance. Statistical power was not calculated if the meta-analysis was significant.

ies. For rectal cancer, these estimates were 1.37 (95% CI: 0.98, 1.92) for the five mortality studies and 1.42 (95% CI: 0.70, 2.88) for the lone incidence study. This stratified analysis suggests that incidence studies in our analysis do not yield different relative risk estimates from mortality studies.

Discussion

The data support a significant association between bladder cancer and exposure to chlorination by-products in drinking water. This association appears to follow a dose-response relationship. Controlling for potential confounders (i.e., sex, smoking, urban living, and occupational risk) does not diminish this association. The association of rectal cancer with chlorination by-products follows a pattern similar to that of bladder cancer.

The association of chlorination by-products with bladder and colon cancer increases with improved exposure assessment and higher overall study quality. This pattern strengthens the evidence for an association between both rectal and bladder cancer and chlorination by-products and suggests that the true relative risks may be higher than those listed here.

Chlorination by-products are associated with cancer of the rectum, but not cancer of the colon. The bladder and the rectum both serve a similar physiological function by storing concentrated excretory products. One might speculate that the epithelial tissue at both sites is exposed to higher levels of chlorination by-prod-

TABLE 5—Relative Risk Estimates for the Association of Exposure to Chlorination By-products with Cancer of the Bladder and Colon, from Studies Grouped by Sex, Adjustment for Confounders, and Quality Scores

Site	Category	Group	Relative Risk Estimate	95% CI
Bladder	All relevant studies		1.21	1.09, 1.34
	Sex	Females	1.17	1.03, 1.34
		Males	1.24	0.97, 1.57
		Combined[a]	1.22	1.06, 1.37
	Confounders	Population density	1.20	1.08, 1.33
		Smoking	1.25	1.09, 1.45
		Occupation	1.22	1.06, 1.41
	Overall quality score	High	1.25	1.09, 1.45
		Low	1.16	0.96, 1.40
	Exposure quality score	High	1.25	1.09, 1.45
		Low	1.16	0.96, 1.40
Rectum	All relevant studies		1.38	1.01, 1.87
	Sex	Females	1.10	0.90, 1.36
		Males	1.24	0.86, 1.79
		Combined[a]	1.42	0.99, 2.03
	Confounders	Population density	1.33	1.02, 1.73
	Overall quality score	High	1.91	1.56, 2.35
		Low	1.07	0.89, 1.29
	Exposure quality score	High	1.57	1.10, 2.25
		Low	1.23	0.89, 1.70
Colon	All relevant studies		1.11	0.91, 1.37
	Sex	Females	1.19	0.93, 1.53
		Males	1.09	0.81, 1.48
		Combined[a]	1.02	0.86, 1.22
	Confounders	Population density	1.09	0.89, 1.32
	Overall quality score	High	1.10	0.79, 1.53
		Low	1.13	0.86, 1.48
	Exposure quality score	High	1.05	0.72, 1.54
		Low	1.11	0.87, 1.41

[a]This group includes only those studies that considered both sexes. Hence the results are different from the values for all relevant studies.

TABLE 6—Meta-analytic Estimates of the Association between Bladder and Colon Cancer at Specified Levels of Cumulative Exposure to Chlorination By-products

Site	Level of Exposure	Relative Risk Estimate	95% CI
Bladder	Low	1.03	0.85, 1.24
	Moderate[a]	1.20	1.05, 1.37
	Moderate[b]	1.20	1.04, 1.38
	High	1.41	1.24, 1.62
Rectum	Low	1.13	0.61, 2.09
	Moderate[a]	1.29	1.00, 1.67
	Moderate[b]	1.49	1.10, 2.01
	High	2.04	1.18, 3.53

[a]This category includes only data from those studies that ranked a middle level of exposure.
[b]This category includes all studies for which exposure was dichotomous, in addition to studies ranked a middle level of exposure.

ucts and is therefore at increased risk for the development of neoplasia.

The most important potential confounder not adjusted for in these studies is diet. If consumption of a diet high in fat and low in fiber were associated with consumption of chlorinated water, one could argue that the observed association is simply a surrogate for dietary risk. We know of no such association. In addition, if diet were responsible for the observed association with rectal cancer, we would expect to see the same association with colon cancer. The marked differences in these associations tend to refute the contention that diet is an explanatory factor for observed associations.

With regard to other sites for neoplastic disease, relative risks ≤ 1.2 cannot be excluded with a high degree of confidence for any of the sites considered. A relative risk ≥ 1.6 for cancer associated with chlorination by-products can, however, be excluded with a high degree of confidence, on the basis of current evidence, for most sites. The meta-analysis for brain cancer lacks the power to exclude with any certainty even a high relative risk for cancer.

The present study has identified a clear and significant association between neoplastic disease and the consumption of water containing chlorination by-products. Precise cause and effect cannot be determined. The possibility that chlorination may serve as a marker for some other aspect of drinking water quality or an associated geographic or demographic variable cannot be entirely discounted, given the available research. If chlorination by-products are carcinogens, the identity of the compound or compounds involved is unclear. One agent actively being investigated in this context is a chlorinated hydroxy-furanone recently identified as an extremely potent mutagen present in chlorinated surface water.[30]

If we posit that the association demonstrated between chlorination by-products and cancer represents a causal relationship in some way, the positive relative risks identified for bladder and colon cancer can best be interpreted in the context of attributable proportions (i.e., that portion of cases that can be attributed to a given risk factor[32]). By a conservative estimate, based on the 54% of the population that consumes chlorinated surface water (rather than the 75% that consumes chlorinated water), a 21% increase in the risk for bladder cancer and a 38% increase in the risk for rectal cancer translate into attributable proportions of 9% and 15%, respectively. With US incidence rates for bladder and rectal neoplasms of 47 000 and 44 000 cases per year, respectively,[33] these proportions suggest that about 4200 cases (9%) of bladder cancer per year and 6500 cases (18%) of rectal cancer per year are associated with the consumption of chlorinated water. The higher estimates for relative risk associated with more accurate assessment of exposure suggest that the true attributable proportions may be higher than those stated here.

Of the original papers included in the meta-analysis, only three of the seven bladder cancer studies and two of the six rectal cancer studies were statistically significant. Consequently, a standard review of the literature might conclude, erroneously, that no significant risk exists. Seemingly small risks posed by environmental factors may have major health impacts because of the large numbers of people exposed. The results of the meta-analysis demonstrate the utility of this methodology in the identification of important but elusive relative health risks and suggest that a wider application of meta-analysis is warranted.

Our findings are in no way intended to suggest that the disinfection of drinking water should be abandoned. The potential health risks of microbial contamination of drinking water greatly exceed the risks described above. Nonetheless, these findings should provide an impetus to identify, develop, and implement disinfection strategies that are not associated with adverse health effects. □

References

1. White CW. *Handbook of Chlorination.* New York: Van Nostrand Reinhold; 1972: 278–295.
2. Cantor KP. Epidemiologic studies of chlorination by-products in drinking water: an overview. In: Jolley RL, ed. *Water Chlorination: Environmental Impact and Health Effects* 4th ed. Ann Arbor, Mich: Ann Arbor Scientific Publishers; 1983: 1381–1397.
3. Rook JJ. Formation of haloforms during chlorination of natural waters. *J Soc Water Treatment Examination.* 1974;23:234–243.
4. Symons JM, Bellar TA, Carswell JK, et al. National Organics Reconnaissance Survey for Halogenated Organics. *J Am Water Works Assoc.* 1975;67:634–646.
5. Page T, Harris RH, Epstein SS. Drinking water and cancer mortality in Louisiana. *Science.* 1976;193:55–57.
6. Kuzma RJ, Kuzma CM, Buncher CR. Ohio drinking water source and cancer rates. *Am J Public Health.* 1977;67:725–729.
7. Cantor KP, Hoover R, Mason TJ, McCabe LJ. Association of cancer mortality with halomethanes in drinking water. *J Natl Cancer Inst.* 1978;61:979–985.
8. Hogan MD, Chi PY, Hoel DG, Mitchell TJ. Association between chloroform levels in finished drinking water supplies and cancer mortality rates. *J Environ Pathol Toxicol.* 1979;2:873–887.
9. Carlo GI, Mettlin CJ. Cancer incidence and trihalomethane concentrations in a public drinking water system. *Am J Public Health.* 1980;70:523–525.
10. Bean JA, Isacson P, Hausler WJ Jr, Kohler J. Drinking water and cancer incidence in Iowa: I. trends and incidence by source of drinking water and size of municipality. *Am J Epidemiol.* 1982;116:912–923.
11. Sullivan JW. Epidemiologic survey of bladder cancer in greater New Orleans. *J Urol.* 1982;128:281–283.
12. Beresford SAA. Cancer incidence and reuse of drinking water. *Am J Epidemiol.* 1983;117:258–268.
13. Carpenter LM, Beresford SAA. Cancer mortality and type of water source: findings from a study in the UK. *Int J Epidemiol.* 1986;15:312–319.
14. Marienfeld CJ, Collins M, Wright H, Reddy R, Shoop G, Rust P. Cancer mortality and the method of chlorination of public drinking water: St. Louis City and St. Louis County, Missouri. *JEPTO.* 1986; 7:41–158.

15. Alvanja M, Goldstein I, Susser M. A case-control study of gastrointestinal and urinary tract cancer mortality and drinking water chlorination. In: Jolley RL, Gorchev H, Hamilton DH Jr, eds. *Water Chlorination: Environmental Impact and Health Effects.* 2nd ed. Ann Arbor, Mich: Ann Arbor Science Publishers; 1978:395–409.
16. Brenniman GR, Vasilomanolakis-Lagos J, Amsel J, Tsukasa N, Wolff AH. Case-control study of cancer deaths in Illinois communities served by chlorinated or non-chlorinated water. In: Jolley RL, Brungs WA, Cumming RB, et al., ed. *Water Chlorination: Environmental Impact and Health Effects.* 3rd ed. Ann Arbor, Mich: Ann Arbor Scientific Publishers; 1980: 1043–1057.
17. Young TB, Kanarek MS, Tsiatis AA. Epidemiologic study of drinking water chlorination and Wisconsin female cancer mortality. *J Natl Cancer Inst.* 1981;67:1191–1198.
18. Gottlieb MS, Carr JK, Clarkson JR. Drinking water and cancer in Louisiana: a retrospective mortality study. *Am J Epidemiol.* 1982;116:652–667.
19. Sullivan JW. Epidemiologic survey of bladder cancer in greater New Orleans. *J Urol.* 1982;128:281–283.
20. Lawrence CE, Taylor PR, Trock BJ, Reilly AA. Trihalomethanes in drinking water and human colorectal cancer. *J Natl Cancer Inst.* 1984;72:563–568.
21. Cragle DL, Shy CM, Struba RJ, Siff EJ. A case-control study of colon cancer and water chlorination in North Carolina. In: Jolley RL, ed. *Water Chlorination Chemistry, Environmental Impact and Health Effects.* Chelsea, Mich: Lewis Publishers; 1985: 153–159.
22. Zierler S, Danley RA, Feingold L. Type of disinfectant in drinking water and patterns of mortality in Massachusetts. *Environ Health Perspect.* 1986;69:275–279.
23. Cantor KP, Hoover R, Hartge P, et al. Bladder cancer, drinking water source, and tap water consumption: a case-control study. *J Natl Cancer Inst.* 1987;79:1269–1279.
24. Young TB, Wolf DA, Kanarek MS. Case-control study of colon cancer and drinking water trihalomethanes in Wisconsin. *Int J Epidemiol.* 1987;16:190–197.
25. Wilkins JR, Comstock GW. Source of drinking water at home and site-specific cancer incidence in Washington County, Maryland. *Am J Epidemiol.* 1981;114:178–190.
26. Schulte PA, Ringen K, Hemstreet GP, et al. Risk factors for bladder cancer in cohort exposed to aromatic amines. *Cancer.* 1986; 58:2156–2162.
27. Zierler S, Fiengold L, Danley RA, Craun G. Bladder cancer in Massachusetts related to chlorinated and chloraminated drinking water: a case-control study. *Arch Environ Health.* 1988;43:195–200.
28. Chalmers TC, Smith H Jr, Blackburn B, et al. A method for assessing the quality of a randomized control trial. *Controlled Clin Trials.* 1981;2:31–49.
29. Longnecker MP, Berlin JA, Orza MJ, Chalmers TC. A meta-analysis of alcohol consumption in relation to risk of breast cancer. *JAMA.* 1988;260:652–656.
30. Kronberg L, Holmboom B, Tikkanen L. Identification of the strong mutagen 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone and its geometric isomer E-2-chloro-3-(dichloromethyl)-4-oxybutenoic acid in mutagenic fractions of chlorine-treated humic water and in drinking water. In: Jolley RL, Condie LW, Johnson JD, et al., ed. *Water Chlorination Chemistry, Environmental Impact and Health Effects.* Chelsea, Mich: Lewis Publishers; 1987:137–146.
31. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clin Trials.* 1986; 7:177–188.
32. Rothman K. *Modern Epidemiology.* Boston, Mass: Little Brown & Co; 1986:284–310.
33. *Cancer Facts and Figures.* Atlanta, Ga: American Cancer Society; 1989.

Call for Abstracts for Injury Control/Emergency Services Late-Breaker Session

The Injury Control and Emergency Health Services Section will again sponsor a "late-breaker" session during the APHA 1992 annual meeting in Washington, DC. The session will be held on Tuesday, November 10, and will feature work completed within the last few months (after the deadline for consideration in the regular symposia of the APHA annual meeting).

Submit abstracts of 250 words or fewer (any format) and a return envelope to Richard Wanweiler, Division of Injury Control, Centers for Disease Control, Mail Stop F-36, Atlanta, GA 30333; tel. (404) 488-4680.

Abstracts must be received by September 1, 1992. Decisions will be made by September 15, 1992.

Epidemiological Data on US Coal Miners' Pneumoconiosis, 1960 to 1988

Michael D. Attfield, PhD, and Robert M. Castellan, MD, MPH

ABSTRACT

Objectives. Statistics on prevalence of pneumoconiosis among working underground coal miners based on epidemiologic data collected between 1960 and 1988 are presented. The main intent was to examine the time-related trend in prevalence, particularly after 1969, when substantially lower dust levels were mandated by federal act.

Methods. Data from studies undertaken between 1960 and 1968 were collected and compared. Information for the period 1969 to 1988 was extracted from a large ongoing national epidemiologic study. Tenure-specific prevalence rates and summary statistics derived from the latter data for four consecutive time intervals within the 19-year period were calculated and compared.

Results. The results indicate a reduction in pneumoconiosis over time. The trend is similar to that seen in a large radiologic surveillance program of underground miners operated concurrently.

Conclusions. Although such factors as x-ray reader variation, changes in x-ray standards, and worker self-selection for examination may have influenced the findings to some extent, adjusted summary rates reveal a reduction in prevalence concurrent with reductions in coal mine dust levels mandated by federal act in 1969. (*Am J Public Health.* 1992; 82:964–970)

Introduction

The National Institute for Occupational Safety and Health (NIOSH) receives frequent requests for information on the prevalence of coal workers' pneumoconiosis (CWP), but no current and comprehensive statistics are available in a convenient format. Two reports have been compiled to fill this gap. The first, this paper, concerns the findings from epidemiological studies undertaken from 1960 to 1988, including the National Study of Coal Workers' Pneumoconiosis (NSCWP), a large nationwide epidemiological study managed by NIOSH. The second[1] deals with information from the Coal Workers' X-ray Surveillance Program (CWXSP) from 1970 to 1986.

Among the provisions of the 1969 Federal Coal Mine Health and Safety Act[2] was a directive that a research program be set up to inquire into the respiratory health of US coal miners. The NSCWP, begun in 1969 and still in progress, was the response to this mandate. Based on mines selected from all the major coalfields, it is the largest and longest-running epidemiological study of coal miners in the United States. Many of the mines were included in some earlier medical[3] and environmental[4] investigations. Medical examinations were undertaken at these NSCWP mines up to four times between 1970 and 1988.

The NSCWP was designed to estimate levels of lung disease in underground coal miners, and to study the relationship between disease and the level and constituents of dust exposure. Results from the study have shed a great deal of light on CWP in US coal miners,[5–10] as well as on other lung diseases.[11–13] To date, however, the CWP reports based on NSCWP data have been tailored to suit a variety of specific inquiries, and their resulting lack of uniformity in format makes comparison among them difficult. Furthermore, careful review of these reports reveals apparent inconsistencies between the prevalences of CWP published at different times—notably, between those published at the start of the study and those disseminated later. The objective of the present paper is to present all the NSCWP-generated CWP prevalence estimates, along with those of preceding US studies, in a unified manner, and to review and explain the differences and inconsistencies.

Coal Workers' Pneumoconiosis. CWP is almost always identified in living coal miners by examination of the chest radiograph. Coal mine dust deposits in the lung give rise to radiographic opacities. Small (less than 1 cm diameter) opacities are classified into four major categories (0, 1, 2, and 3) by reference to standard films representing levels of profusion. The current reference is the 1980 International Labour Office (ILO) classification system.[14] Small opacities are also classified according to shape (rounded or irregular). The actual approach to differentiating between small rounded and small irregular opacities and to determining profusion has differed over the years (see below). This report concentrates on small rounded opacities because that type has been conventionally associated with coal mine dust exposure.[15]

Simple CWP is manifested by small opacities on the chest radiograph. It is dis-

The authors are with the Division of Respiratory Disease Studies at the National Institute for Occupational Safety and Health in Morgantown, WV.

Requests for reprints should be sent to Michael D. Attfield, PhD, MS 234, ALOSH, 944 Chestnut Ridge Road, Morgantown, WV 26505.

This paper was submitted to the Journal July 1, 1991, and accepted with revisions March 23, 1992.

tinguished from complicated CWP, also known as progressive massive fibrosis (PMF), a more advanced form of CWP manifested by large radiographic opacities (at least 1 cm in diameter) on the chest radiograph.

Historical Information. Prior to the 1969 promulgation of federal limits on airborne respirable dust concentration in US underground coal mines, respirable dust levels substantially exceeded the current federal limit of 2 mg/m^3. For example, Jacobson[4] showed that average dust concentrations for certain miners working at the coal face were close to 6 mg/m^3, and that the mean concentrations for most other underground jobs were above 2 mg/m^3.

Shown in Table 1 are summary results on prevalence of category 1 or greater CWP derived from eight reports published between 1961 and 1970.[3,16-22] Eleven regions were involved, ranging from the anthracite coalfield of Pennsylvania to southern mines located in Alabama and to western mines in Utah. This review extends that of Morgan.[23] With the exceptions noted in the footnotes to the table, all the information concerns working coal miners and is based on the 1959 ILO classification.[24] For these tabulations, category Z opacities (suspicious markings) were classified with those in category 0.

Table 1 arranges the studies according to rank of coal (higher coal rank implies an older coal of higher percentage carbon), there being evidence that CWP incidence varies with coal rank.[5] The regions considered were high rank (anthracite), medium-high rank (low volatile bituminous), medium rank (mainly high volatile "A" bituminous), and medium-low rank (mainly high volatile types "B" and "C" bituminous). The results confirm the previously reported association between CWP prevalence and coal rank. Mean mining tenure and age were similar across coal rank groups and therefore were unlikely to have been responsible for the observed coal rank trend. A fair amount of interstudy variation within coal rank groups is seen, which may have resulted from the use of different x-ray standards and readers. Varying rates of worker participation and differences in dust exposure (data not available) may also have contributed to the variation.

Tenure-specific prevalences from these studies, where available, are plotted in Figure 1. In every study, prevalence of small rounded opacities increased with increasing tenure. Agreement is quite good for those studies from the lower rank regions.

Information on more advanced profusions of small opacities was available from only four studies, and findings for category 2 or greater small opacities varied from 22% in working miners in the high rank region,[16] to 21% and 33% in medium-high rank coal miners,[18,19] to 7% in medium rank miners.[22] Large opacity prevalence also varied fairly consistently with

TABLE 1—Prevalence Rates of Coal Workers' Pneumoconiosis (Category 1 or Greater) for Some US Studies Undertaken between 1961 and 1970, In Order of Coal Rank

Study (by Probable Coal Rank)	Prevalence of CWP (Category 1 or Greater)	PMF	Mean Age, y	Mean Tenure, y		Comments (See Below)
				Mine	UG	
High rank						
A. Eastern Pennsylvania[16]	22	9.6	55	26	...	1
B. Eastern Pennsylvania[17]	34	...	45	...	22	1, 2
Medium-high rank						
C. Central Pennsylvania[18]	25	8.3	47	27	...	1
D. Southern West Virginia[19]	46	7.2	52	...	25	1, 3
E. Southern West Virginia[20]	14	5.4	44	...	...	4
Medium rank						
F. Appalachia[3]	10	3.0	47	...	22	5
G. Eastern West Virginia[20]	6	1.1	43	...	...	4
H. Northern West Virginia[21]	7	0.9	48	...	19	1, 3
I. Western Pennsylvania[22]	9	3.7	47	26	...	1
Medium-low rank						
J. Illinois/Indiana[3]	6	1.5	48	...	20	5
K. Utah[3]	4	0.7	51	...	20	5

Notes. PMF = progressive massive fibrosis; Mine = all work in coal mining; UG = work underground.
1. The 1959 ILO classification[24] was used.
2. Exact figure for average age is not given; figure given here has been estimated from age distribution data.
3. Group contains both current and ex-miners, and could not be subdivided.
4. Radiographic classification is not stated but is probably the ILO 1959 scheme,[24] given that the study was undertaken from 1963 to 1964.
5. The classification used is not stated explicitly. The reading sheet shown in the report looks very similar to that of the ILO 1968 classification,[25] but Morgan[23] states that the ILO 1959 classification system was used.

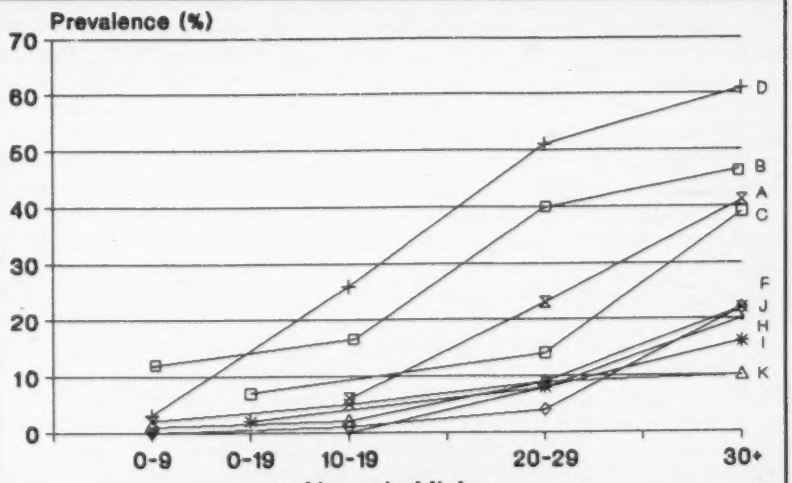


FIGURE 1—Prevalence of category 1 or greater small rounded opacities by tenure in mining for studies of underground US coal miners published between 1961 and 1970. (See Table 1 for the studies corresponding to the letters.)

coal rank, from about 10% in the high rank region[16] to about 1% in the medium-low rank region[3] (Table 1). Prevalences of large opacities were markedly higher in workers who were close to retirement age, with 19%, 18%, and 9% prevalences being reported for high, medium-high, and medium rank regions of Pennsylvania, respectively.[16,18,22]

NSCWP Methods

Survey Information

The NSCWP has been conducted in rounds of medical surveys. Round 1 examinations were conducted from 1969 to 1971, round 2 from 1972 to 1975, round 3 from 1977 to 1981, and round 4 from 1985 to 1988. The procedure in each of the first three rounds was to examine the miners at the mine site, either before or after their work shift, using a mobile x-ray trailer. Forty mines have been visited at one time or another over the course of the study. Thirty-one mines were visited at round 1. The principal criteria at round 1 for the selection of mines were (1) an expected life of the mine of at least 10 years, (2) a work force of at least 100 miners, (3) a geographic and geologic spread, and (4) accessibility for the field examination trailer. Owing to unexpected mine closings, nine other mines were substituted in round 2 following the same selection criteria.

Rounds 1 to 3 were a series of cross-sectional studies at virtually the same group of mines. Round 4 was organized very differently, being a follow-up of selected miners from the first two rounds. As a consequence, most of the round 4 examinations were not undertaken at the original mine sites but comprised what were termed "community surveys." Based on residence addresses, suitable clusters of miners and ex-miners were identified from the follow-up cohort. The mobile examination trailer was set up at a central location in each of 22 communities where clusters occurred, and individual contacts were made to encourage miners in the follow-up cohort to participate in the examination. In addition, mine surveys were held at three of the original mines; more were not possible because many of the original mines had closed or were nonoperational owing to the coal mining recession that occurred in the mid-1980s.

Low participation rates have been a problem with some rounds of the NSCWP. Although participation was very good at round 1 (accounting for 90% of those eligible at time of examination), the next two surveys were marked by a substantial decline in participation to 75% at round 2 and to 52% at round 3. The change in study design and survey strategy in round 4 led to an improvement in participation, accounting for 70% of those selected at the community studies and better rates at the mine site surveys.

X-ray Readers, X-ray Classifications, and Associated Methods

Apart from a switch to rare earth screens in 1985, the x-ray equipment and methods have changed little over the course of the NSCWP. Throughout the study, three x-ray readings were obtained for each chest radiograph. To date, resources have not been available to reread all the films using the same readers, protocol, and classification. Therefore, the information presented here is based on the original readings made at each round of the NSCWP.

The round 1 films were classified using the Union Internationale Contre le Cancer (UICC)/Cincinnati (U/C) classification system.[26] Of the three interpreters employed, only two specifically classified small rounded opacities (readers 1 and 2). These were both expert readers, who had collaborated with others in developing the UICC/Cincinnati classification system.[26] Later examination of their readings revealed, however, that reader 1 found improbably high levels of abnormality. His results give rise to prevalences that deviate substantially from every other source, including prior studies, other NSCWP surveys, concurrent surveillance program data,[1] and readings from a group of British interpreters.[27] Because of this discrepancy, his results have not been included in this analysis. Instead, the round 1 data presented here are those of the single remaining reader who classified small rounded opacities (reader 2); this interpreter's findings were much more consistent with those of the other sources noted above. (The readings of reader 1 were included in an earlier analysis,[5] which thereby reported a considerably greater prevalence of CWP than that given here.)

The 1971 ILO U/C classification system[28] was adopted for round 2. Readers 2 and 3 from round 1, along with seven other B readers (readers who have taken and passed a NIOSH-administered test of competency in reading for the pneumoconioses),[29] classified radiographs in the second round. All cases included in this analysis were based on median determinations of three readings of small rounded opacities.

The round 3 results presented here were also based on median determinations of small rounded opacities, this time made by three B readers using the 1971 classification. None of the three readers had participated in the earlier rounds.

For round 4, each film was independently read by three B readers using the 1980 ILO classification system.[14] Because the 1980 system does not permit profusions of small rounded and small irregular opacities to be classified separately, small rounded opacity profusion was put equal to the overall profusion score if rounded opacities were reported to predominate. However, if irregular opacities were said to be more common, the small rounded opacity profusion was set to 0. This approach probably tends to exaggerate the rounded profusion category when the primary type is rounded, and to underestimate the rounded profusion category when irregular opacities predominate.

To select readers for round 4, readings of all B readers who had read films for round 4 of the CWXSP were tabulated, and the readers were ranked according to amount of abnormality detected. Three readers lying in the midrange were then selected. None of these readers had participated in earlier rounds of the NSCWP.

In addition to small rounded opacities, this report also presents information on progressive massive fibrosis. The data were derived from reader 2 at round 1 and from medians of three readings at each of the subsequent rounds. Some summary statistics are also given on small combined opacity prevalences obtained similarly from reader 2 at round 1 and from medians elsewhere.

Throughout the NSCWP, occupational histories were obtained from the miners using a standardized questionnaire that varied little over the four rounds.

The tabulations that follow are based on all complete sets of data (readable x-rays and available tenure information) from each round, with the exception of round 4. For round 4, ex-miners who said they had left work because of health reasons were not included in order to make round 4 results more compatible with those from the earlier rounds. In general, these ex-miners had more abnormality than the others. However, laid-off miners and ex-miners who had taken other jobs were included with the current miners in the round 4 tabulations. The great major-

ity of these miners had left mining within a year or two before the medical survey.

NSCWP Results

Distribution of Examined Miners by Geologic Region

Table 2 shows how the data for those cases with valid x-ray readings were distributed across four geologic regions. A slight temporal movement from high to low rank coal areas is apparent. This generally reflects the pattern of mine closings that has occurred over the past two decades, particularly in the high rank region. Note, however, that the percentages of miners in the medium and medium-low rank regions remain almost constant over the four rounds.

TABLE 2—Miner Location by Round and Coal Rank Region as Defined by the National Study of Coal Workers' Pneumoconiosis

Region	Round 1[a] Miners	Round 2[a] Miners	Round 3[a] Miners	Round 4[a] Miners
High rank, %	6	3	0	0
Medium-high rank, %	15	14	19	17
Medium rank, %	55	58	55	51
Medium-low rank,%	25	25	27	32
Number of participants	9078	9347[b]	5275	3534
Number of x-ray readings	9034	8385[b]	5131	3451

[a]Percentages were based on the number of x-ray readings.
[b]Films for miners at two mines at round 2 were read by one reader only, leading to a larger difference between the numbers of participants and readings than exist for the other three rounds.

Distribution of Examined Miners by Tenure

The tenure distribution of coal miners has changed considerably over time owing to economic activity and other factors (Table 3). Prior to 1973, the coal mining work force was quite elderly, as the industry had been undergoing a long-term decline. Over the next few years, many of the older and longer-tenured miners left mining and were replaced by younger workers. More importantly, a large number of new miners were hired in the wake of the energy crisis–related mining expansion that took place in the mid-1970s. Since then, the mean age of the work force has been gradually rising.

TABLE 3—Tenure in Mining by Round as Defined by the National Study of Coal Workers' Pneumoconiosis

Tenure Group	% Round 1 Miners (n = 9034)	% Round 2 Miners (n = 8385)	% Round 3 Miners (n = 5131)	% Round 4 Miners (n = 3451)	% 1986 Survey Miners (n = 78 741)
0	0.0	2.4	1.1	2.4	2.0
1	10.2	7.4	3.5	1.0	2.0
2–4	10.0	19.9	18.2	2.9	8.9
5–9	7.6	13.2	32.3	12.4	29.2
10–14	5.6	5.8	13.9	26.7	35.0
15–19	8.0	6.0	5.1	29.1	14.1
20–24	14.9	8.5	4.0	9.8	3.6
25–29	16.1	14.4	4.8	5.0	1.6
30+	27.5	22.3	17.1	10.7	3.7

Notes. Data for rounds 1 to 3 reflect the mining work force tenure distributions except for the influence of selective participation. Round 4 was not a cross-sectional study but largely a follow-up group, and the tenure distribution reflects that. The 1986 survey data are projected statistics for coal miners underground and on the surface at underground mines in the bituminous and anthracite areas. They were derived from a Bureau of Mines mail survey of coal mine operators.[30]

Also included in Table 3 is a tenure distribution projected for the underground coal mining work force for 1986. This was derived from data collected from a mail survey of coal mine operators undertaken by the Bureau of Mines.[30] The data, which apply to both underground miners and surface workers at underground facilities in the bituminous and anthracite regions, were supplied on computer tape by the bureau. Because these data were rounded up to the nearest year, one year was subtracted to make them consistent with the radiological data, for which rounding down was performed. Based on the bureau's estimates, the round 4 tenure distribution clearly does not represent the general underground mining work force at that time. This is not surprising, given that round 4 participants represent a cohort followed from rounds 1 and 2.

Small Rounded Opacities

In every round of the NSCWP, a clear trend of increasing prevalence of CWP with tenure is evident (Figures 2 and 3). Round 4 tenure-specific rates of category 1 or greater small rounded opacities were about one-third those of round 1 (Figure 2), while rates of category 2 or greater opacities showed larger drops from round 1 to round 4 (Figure 3).

Large Opacities

Figure 4 provides information on the prevalence of progressive massive fibrosis. These data for large opacity prevalence are probably more prone to selection effects than are the data for small opacities, given that miners with severe CWP are more likely to leave mining for health reasons (and with compensation provided by the Federal Coal Mine Health and Safety Act ["Black Lung Act"] of 1969). There is a trend toward lower prevalences from round 1 through round 3, which probably reflects the effect of both lower dust exposures and loss of miners with progressive massive fibrosis from the work force. The prevalences at round 4 are similar to those at round 3. The reason for this is not clear although the increased participation rates at round 4 may be responsible. The round 4 data are currently under more detailed study, including an evaluation of 15-year incidence.

Standardized Rates

For each round, estimated summary statistics for small rounded opacity prevalence, obtained from the Bureau of Mines survey data[30] and standardized to the common tenure distribution, were derived for small rounded opacities of category 1 or greater and of category 2 or greater. These are shown in Table 4. They indicate a slight drop in prevalence from round 1 to round 2, an approximate halving in prevalence from round 2 to round 3, and a further drop from round 3 to round 4.

Small Combined Opacities

Although small rounded opacities have been more commonly associated

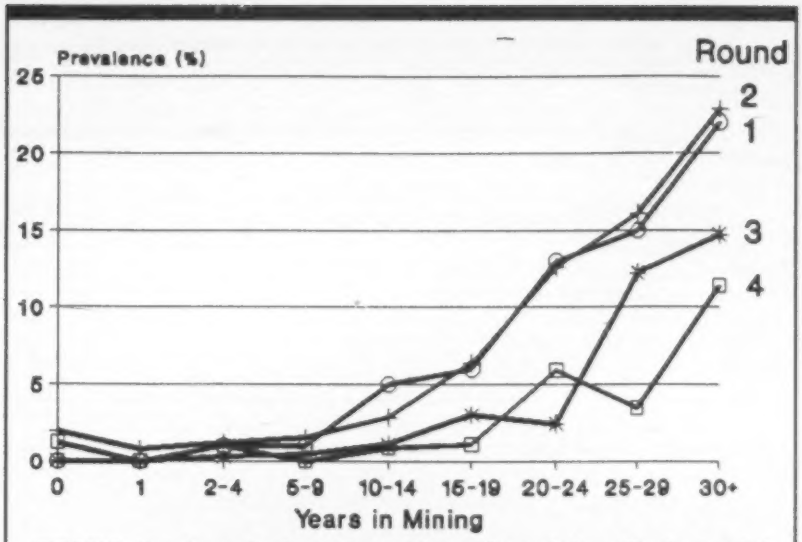


FIGURE 2—Prevalence of category 1 or greater small rounded opacities by round of the National Study of Coal Workers' Pneumoconiosis and by tenure group. Single reader at Round 1; median readings for other rounds.

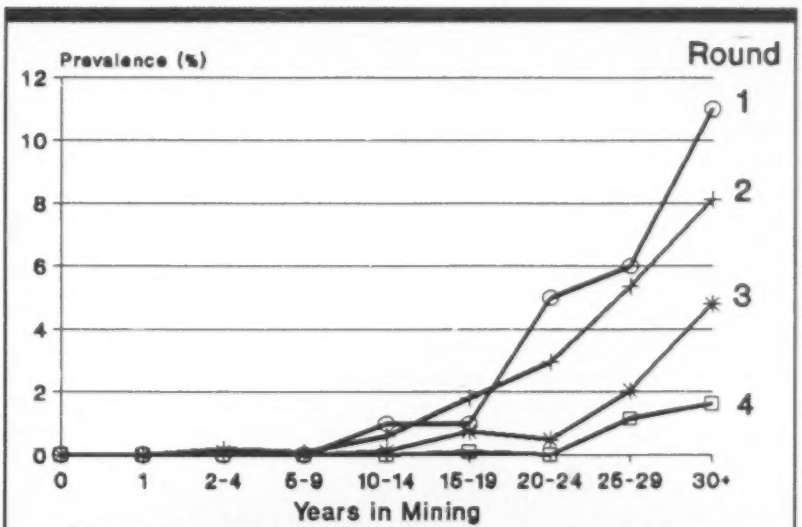


FIGURE 3—Prevalence of category 2 or greater small rounded opacities by round of the National Study of Coal Workers' Pneumoconiosis and by tenure group. Single reader at Round 1; median readings for other rounds.

with CWP, there is some evidence that small irregular opacities may also be associated with coal mine dust exposure.[31] For this reason, and because different readers sometimes report similar profusions yet differ on the predominant type of small opacity on chest radiographs, summary adjusted statistics for combined opacities are also shown in Table 4. These prevalences are somewhat higher, as expected, but do parallel the downward trends seen for small rounded opacities.

Discussion

The prevalences of CWP shown in Figures 2 to 4 and summarized in Table 4 reveal clear downward trends over time. Similar trends are evident in data from a large federal x-ray surveillance program for underground coal miners.[1] Both sets of data suggest that prevalences at round 4 were about a third of their levels at round 1. Reductions in reported prevalence of more severe categories of pneumoconiosis (e.g., category 2 or greater) were somewhat greater.

The observed trends occurred concurrently with reductions in dust exposure mandated by the 1969 Federal Coal Mine Health and Safety Act and recorded by the Mine Safety and Health Administration. According to the report by Parobeck and Jankowski,[32] respirable dust levels in US mines dropped from averages of about 6 mg/m^3 for face jobs in 1968 to 1969 to less than 1.5 mg/m^3 10 years later. Although there is evidence that the actual decline in dust levels may not have been as great as reported,[33,34] it is very likely that most underground coal miners experience lower exposures today than they did prior to passage of the 1969 Federal Coal Mine Health and Safety Act.

Although reductions in dust levels and other CWP prevention measures are the most likely reason for the declines in prevalence of radiological abnormality, there may be some degree of confounding with other effects, such as selective participation and exodus of miners with lung disease from the work force. Artifactual effects associated with x-ray reading, such as the use of different readers and international x-ray classifications systems, and possible changes in readers' perceptions over time may also be confounded with real trends in prevalence. It is difficult to assess how these factors may have influenced the results presented here. To obtain a clearer understanding of temporal trends in the prevalence and severity of CWP in coal miners, a study is planned that will involve the rereading of large numbers of chest radiographs selected from all four rounds of the NSCWP. This will be designed to control for potential confounding associated with these artifactual effects.

The Bureau of Mines survey data given in Table 3 indicate that the 1986 underground coal mining work force was fairly young, with most miners having worked fewer than 20 years in mining. As a consequence, use of this distribution for standardization leads to rates reflecting prevalence among the less experienced miners. If it is preliminary evidence on the efficacy of the current federal dust limit that is being sought, use of this distribution is very appropriate because, to date, only those with limited tenure have been working under the lower dust limits mandated by the 1969 act. However, use of this distribution does not lead to a fair picture of prevalence levels and changes among older miners. Simple averages of the tenure-specific rates for miners with 20 or more years of work in

mining indicate drops from 16.9% to 6.9% for category 1 or greater small rounded opacities, and from 7.1% to 0.9% for category 2 or greater small rounded opacities. As might be expected, the drops in prevalence (as measured by the ratios of the prevalences) are not so marked as those obtained for the summary rates in Table 4. This is probably because many of the older miners in round 4 had substantial lung dust burdens that were gained from exposures received before 1969.

Although reliance on only one reading from round 1 is a less than desirable strategy, it was considered unavoidable if data from that round were to be included in this report. Overall, it was felt that the round 1 data were sufficiently reliable to merit inclusion. Reasons for this were (1) the general consistency of the round 1 data with prevalence trends seen in the following three rounds; (2) the consistency of the prevalence trends in the present data compared with trends seen over the same period in coal miner surveillance data[1]; and (3) the similarity between the prevalences from round 1 and those from studies conducted prior to round 1 (medium and medium-low rank data in Figure 1).

One unexpected feature of the results in this report pertains to the Figure 2 tenure-specific data, which show the round 2 rates to be similar to those for round 1. A similar effect was evident over the same period elsewhere.[1] Some consider that this effect may be at least partially artifactual[3,5] and associated with differences in the standard reference films for category 1 that accompanied the 1968 and 1971 versions of the ILO classification. This authority and others (private communication) consider that films that would have been classified as category 0 using the 1968 classification (used at round 1) were more likely to be placed in category 1 under the 1971 x-ray classification (used during rounds 2 and 3). As a result, true declines in prevalence would be offset by apparent increases associated with the different ILO classification system used. However, clear declines in prevalence were seen over the same period for category 2 or greater small rounded opacities. Therefore, where no major effect associated with differences in the standard reference films has been suggested, the proposed explanation appears to have some merit.

Because round 4 was designed to be a follow-up study of participants from earlier rounds of the NSCWP, it was organized rather differently from the cross-sectionally based rounds 1 to 3. Hence, it may be incorrect to include its results with those from the other rounds to assess temporal trends in CWP prevalence. However, because ex-miners who left mining for health-related reasons are excluded from the round 4 results, because the analyses presented are adjusted for tenure, and because most of the ex-miners included left work in mining within a year or two of round 4, inclusion of the round 4 results is probably valid. Further support for their inclusion is provided by results from the surveillance program,[1] in which a similar temporal trend was seen for the same period.

The prevalence data given in this report clearly provide information useful for assessing the efficacy of the current federal dust control regulations in preventing occupational lung disease among underground coal miners. Because this information has its limitations, as noted above, an alternative approach to cross-sectional investigation is the longitudinal study, in which x-ray pairs for individual workers taken over suitable intervals of time are read together under controlled conditions using the same readers and classification system. Such studies have already been undertaken[8,9] and indicate an incidence rate of category 1 or greater small rounded opacities of about 0.5% over 10 years among miners with category 0 on initial examination. This finding is consistent with sharply reduced prevalence rates predicted for the future compared with past levels. Additional longitudinal analyses that include both miners and ex-miners are in progress. Among other things, these will provide needed information on workers who may have left coal mining work because of ill health.

Overall, the results presented here reveal a clear downward trend in the prevalence of CWP in US miners. It seems highly likely that this trend represents a real decline in CWP prevalence, particularly because a similar downward trend has been observed over the same period in data from large numbers of underground coal miners participating in a federal x-ray

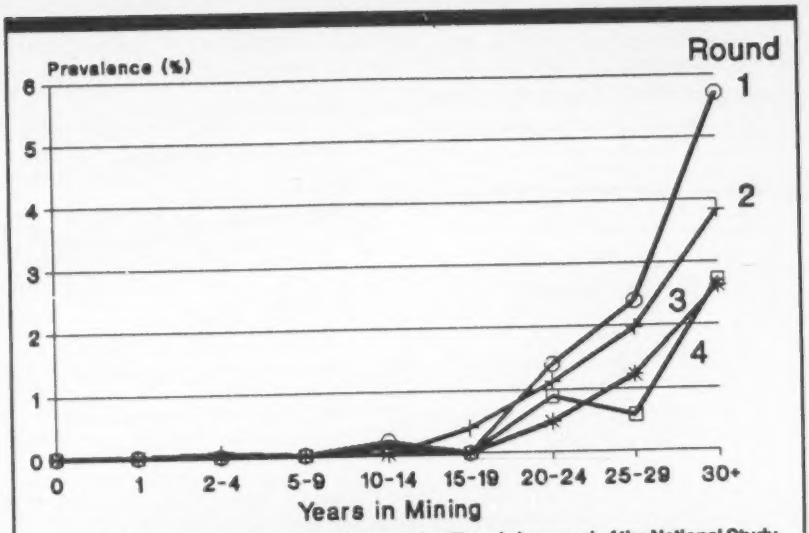


FIGURE 4—Prevalence of progressive massive fibrosis by round of the National Study of Coal Workers' Pneumoconiosis and by tenure group. Single reader at Round 1; median readings for other rounds.

TABLE 4—Summary Coal Workers' Pneumoconiosis Prevalence Statistics from the National Study of Coal Workers' Pneumoconiosis, Based on Small Rounded and Small Combined Opacities, Category 1 or Greater (1+) and Category 2 or Greater (2+), and Adjusted to Common Tenure Distribution

		Adjusted Summary Prevalences, %			
	Category	Round 1	Round 2	Round 3	Round 4
Rounded opacities	1+	4.5	4.1	1.8	1.3
	2+	1.2	1.0	0.4	0.1
Combined opacities	1+	6.6	5.1	3.6	2.3
	2+	1.5	1.2	0.5	0.3

examination program.[1] The findings from this report, coupled with those from longitudinal studies,[8,9] suggest that substantial progress has been made in reducing the prevalence of CWP in underground coal miners. However, owing to the limited periods of study and to other methodological difficulties, these findings should not be taken to imply that coal mining is now risk free with respect to CWP. Furthermore, these results provide no information on dust-induced lung diseases other than pneumoconiosis (e.g., chronic bronchitis and airways obstruction). Continued study of both pneumoconiotic and other lung diseases is necessary to ensure adequate protection of the respiratory health of underground coal miners. □

Acknowledgments

Thanks are due to Barbara Bonnett for help in preparation of the figures and in other programming, and to the many who have helped in preparing and reviewing the manuscript.

References

1. Attfield MD, Althouse RB. Surveillance data on US coal miners' pneumoconiosis, 1970 to 1986. *Am J Public Health.* 1992;82: 971–977.
2. Coal Mine Health and Safety Act. US Pub L No. 91–173, Stat 2917.
3. Lainhart WS. Roentgenographic evidence of coal workers' pneumoconiosis in three areas in the United States. *J Occup Med.* 1969;11:399–408.
4. Jacobson M. Respirable dust in bituminous coal mines in the US. In: Walton WH, ed. *Inhaled Particles III.* Old Woking, England: Unwin Brothers; 1971:745–756.
5. Morgan WKC, Burgess DB, Jacobson G, et al. The prevalence of coal workers' pneumoconiosis in US coal miners. *Arch Environ Health.* 1973;27:221–226.
6. Attfield M, Hudak J. National Coal Study—Prevalence of coal workers' pneumoconiosis: comparison of first and second rounds. In: Rom WN, Archer VE, eds. *Health Implications of New Energy Technologies.* Ann Arbor, Mich: Ann Arbor Science Publishers; 1980:203–212.
7. Amandus HE, Lapp NL, Jacobson G, et al. Significance of irregular small opacities in the radiographs of coal miners in the USA. *Br J Ind Med.* 1976;33:13–17.
8. Attfield MD, Reger RB, Glenn RE. The incidence and progression of pneumoconiosis over nine years in US coal miners: I. principal findings. *Am J Ind Med.* 1984;6: 407–415.
9. Attfield MD, Reger RB, Glenn RE. The incidence and progression of pneumoconiosis over nine years in US coal miners: II. relationship with dust exposure and other potential causative factors. *Am J Ind Med.* 1984;6:417–425.
10. Attfield MD, Morring K. Some investigations into the relationship between coalworkers' pneumoconiosis and dust exposure in US coal miners. *Am Ind Hyg Assoc J.* Accepted for publication.
11. Kibelstis JA, Morgan EJ, Reger R, et al. Prevalence of bronchitis and airway obstruction in American bituminous coal miners. *Am Rev Respir Dis.* 1973;108:886–893.
12. Hankinson JL, Reger RB, Fairman RP, et al. Factors influencing expiratory flow rates in coal miners. In: Walton WH, ed. *Inhaled Particles IV.* Vol 2. Old Woking, England: Unwin Brothers; 1977:737–752.
13. Attfield MD, Hodous TK. Pulmonary function of US coal miners related to dust exposure estimates. *Am Rev Respir Dis.* 1992;145:605–609.
14. International Labour Office. *Guidelines for the Use of ILO International Classification of Radiographs of Pneumoconiosis.* Rev ed. Occupational Safety and Health, Series no. 22 (Rev. 80). Geneva, Switzerland: International Labour Office; 1980.
15. Parkes WR. *Occupational Lung Disorders.* 2nd ed. Butterworths, London: 1982; 197.
16. McBride WW, Pendergrass EG, Lieben J. Pneumoconiosis study of Pennsylvania anthracite miners. *J Occup Med.* 1966;8:365–376.
17. Tokuhata GK, Dessauer P, Pendergrass EP, et al. Pneumoconiosis among anthracite coal miners in Pennsylvania. *Am J Public Health.* 1970;60:441–451.
18. Lieben J, Pendergrass E, McBride WW. Pneumoconiosis study in central Pennsylvania coal mines: I. medical phase. *J Occup Med.* 1961;3:493–506.
19. Hyatt RE, Kistin AD, Mahan TK. Respiratory disease in southern West Virginia coal miners. *Am Rev Respir Dis.* 1964;89: 387–401.
20. Enterline PE. The effects of occupation on chronic respiratory disease. *Arch Environ Health.* 1967;14:189–200.
21. Higgins ITT, Higgins MW, Lockshin MD, et al. Chronic respiratory disease in mining communities in Marion county, West Virginia. *Br J Ind Med.* 1968;25:165–175.
22. McBride WW, Pendergrass E, Lieben J. Pneumoconiosis study of western Pennsylvania bituminous-coal miners. *J Occup Med.* 1963;5:376–388.
23. Morgan WKC. The prevalence of coal workers' pneumoconiosis. *Am Rev Respir Dis.* 1968;98:306–310.
24. International Labour Office. Meeting of experts on the international classification of radiographs of the pneumoconioses. *Occup Health Saf.* 1959:63–69.
25. International Labour Office. *International Classification of Radiographs of Pneumoconiosis.* Revised, 1968. Occupational Safety and Health, Series no. 22. Geneva, Switzerland: International Labour Office; 1970.
26. Bohlig H, Bristol LJ, Cartier PH. UICC/ Cincinnati classification of the radiographic appearances of categories of simple pneumoconiosis: a cooperative study by the UICC committee. *Chest.* 1970;58:57–67.
27. Reger RB, Amandus HE, Morgan WKC. On the diagnosis of coalworkers' pneumoconiosis: Anglo-American disharmony. *Am Rev Respir Dis.* 1973;108:1186–1191.
28. International Labour Office. ILO U/C international classification of radiographs of pneumoconioses 1971. Occupational Safety and Health, Series no. 22 (rev). Geneva, Switzerland: International Labour Office; 1972.
29. Morgan RH. Proficiency examination of physicians for classifying pneumoconiosis chest films. *Am J Radiol.* 1979;132:803–808.
30. Butani SJ, Bartholomew AM. *Characterization of the 1986 Coal Mining Work Force.* Washington, DC: US Bureau of Mines; 1988. Bureau of Mines Information Circular 9192.
31. Collins HPR, Dick JA, Bennett JG, et al. Irregularly shaped small shadows on chest radiographs, dust exposure, and lung function in coalworkers' pneumoconiosis. *Br J Ind Med.* 1988;45:43–55.
32. Parobeck PS, Jankowski RA. Assessment of the respirable dust levels in the nation's underground and surface coal mine operations. *Am Ind Hyg Assoc J.* 1979;40:910–915.
33. Boden LI, Gold M. The accuracy of self-reported regulatory data: the case of coal mine dust. *Am J Ind Med.* 1984;6:427–440.
34. Mine Safety and Health Administration. Peabody Coal Company pleads guilty to dust sample tampering. Washington, DC: Mine Safety and Health Administration; January 17, 1991. Press release USDL 91–23.
35. Merchant JA. Coal workers' pneumoconiosis surveillance in the United States. In: *Proceedings of the Fifth International Conference on Pneumoconiosis, Caracas, Venezuela, 1978.*

Surveillance Data on US Coal Miners' Pneumoconiosis, 1970 to 1986

Michael D. Attfield, PhD, and Rochelle B. Althouse, BS, MS

ABSTRACT

Objectives. Statistics on prevalence of pneumoconiosis among working underground coal miners from data collected as part of a large national radiographic surveillance program between 1970 and 1986 are presented. The main intent was to examine the time-related trend in prevalence over this period, which coincides with historically low dust levels mandated by federal act.

Methods. Tenure-specific prevalence rates and summary statistics derived from them for four consecutive time intervals within the 16-year period were calculated and compared.

Results. The results indicate a reduction in pneumoconiosis over the life of the program. This trend is similar to that seen in epidemiologic studies undertaken concurrently.

Conclusions. Although low participation in the surveillance program and other problems complicate the findings, it appears that reductions in dust exposure mandated by federal act in 1969 have led to lower prevalence of pneumoconiosis among underground coal miners. (*Am J Public Health.* 1992;82:971–977)

Introduction

As part of a lung disease prevention program mandated by the Federal Coal Mine Health and Safety Act,[1] radiographic examinations of the chest are made available to underground coal miners every 5 years. If an x-ray shows certain signs of coal workers' pneumoconiosis (CWP), a miner is given the right to work in a low-dust environment. This procedure, combined with a general control of dust levels mandated by the law, is intended to prevent or slow down the progression of CWP in underground miners and thus reduce the risk of progressive massive fibrosis (PMF).

The administration of these x-rays and of related quality control procedures occurs under the Coal Workers' X-ray Surveillance Program (CWXSP), which is overseen by the National Institute for Occupational Safety and Health (NIOSH). Although primarily serving as a medical monitoring program for underground miners, the CWXSP also provides data that are pertinent to assessing CWP prevalence and to determining the effectiveness of the mandated dust control measures in reducing and eliminating CWP.

This paper provides statistics on the more than 250 000 x-rays collected in the CWXSP since 1970. Information by tenure group is provided, as are various adjusted summary rates reflecting prevalence of CWP among program participants. To permit valid comparisons of data for different periods, details on operational changes are given.

Operation of the CWXSP

Past and present regulations governing the operation of the CWXSP are described in Title 42 of the Code of Federal Regulations.[2–5] In summary, the program is overseen by NIOSH, which, by referring to records on active mines maintained by the Mine Safety and Health Administration, notifies mine operators when its employees are due to be examined. The operator must then file a mine examination plan, which provides information on when and where miners can get their x-rays and on who will do the initial x-ray reading. On receipt of this information, NIOSH determines whether the designated medical facilities and readers are qualified, among other requirements. If a mine operator does not supply a plan, NIOSH prepares one and notifies the operator.

After a plan is approved, it must be posted at the mine site when the time comes due for the examinations. Miners may then present themselves to the designated facility to get an x-ray on their own time and at no cost to themselves. While the miner is at the facility, a brief occupational history is taken using a standard form. The x-ray is classified by the first reader (usually an A reader, but sometimes a B reader [see below]), who typically resides locally. After this, the x-ray film, form, and initial reading must be sent to NIOSH; the x-ray is then sent to a B reader for a further reading. If the two readings are sufficiently similar (definitions of agreement have varied somewhat over time but basically consist of agreement within the same major category of

The authors are with the Division of Respiratory Disease Studies at the National Institute for Occupational Safety and Health in Morgantown, WV.

Requests for reprints should be sent to Michael D. Attfield, PhD, MS 234, ALOSH, 944 Chestnut Ridge Road, Morgantown, WV 26505.

This paper was submitted to the Journal July 1, 1991, and accepted with revisions March 20, 1992.

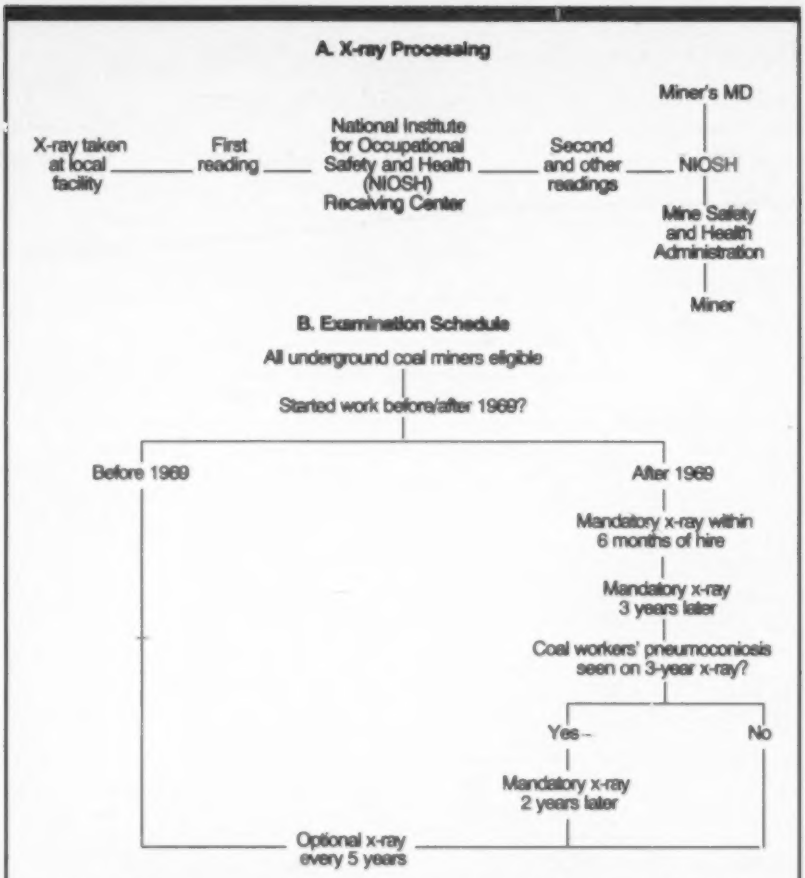


FIGURE 1—Broad processing scheme for x-rays in the Coal Workers' X-ray Surveillance Program.

small opacity profusion), the "final determination" is set equal to the maximum profusion category of the two. If there is greater disagreement than this, further readings are made until sufficient agreement is obtained. The miner is awarded the right to work in a job with a dust exposure of 1 mg/m^3 or less if the final determination is category 1/0 or greater. All program participants have the opportunity to have the results of their x-ray readings (negative or positive) sent to their personal physician. See Figure 1A for details of the x-ray processing.

As a part of a quality control program, the x-ray readers who participate must satisfy certain tests of their competence. Currently, both A and B readers are certified by NIOSH. To become an A reader, the applicant must either attend a special NIOSH seminar or submit six correctly classified chest x-rays. B readers must have passed a special test administered by NIOSH and must undergo regular reexamination to maintain certification.[6] Hundreds of readers have been involved in the program since its inception. As a result, the statistics given in this report should not, in the main, be influenced unduly by individual idiosyncrasies among the readers.

All coal miners working at underground mines are eligible for the periodic x-rays. Currently, miner participation is optional except for new miners, who must have an x-ray within 6 months of starting work at an underground mine and another x-ray 3 years later. A third x-ray, to be taken 2 years later, is mandatory if the second shows signs of CWP. Figure 1B shows the examination schedule schematically.

Between August 1970 and August 1981, the CWXSP was organized into three so-called rounds. Each round was made up of two parts: an initial period when all active mines had to complete examination plans, and a follow-up period when x-rays—mostly those for new miners and for others receiving mandatory x-rays—were being taken. However, logistical problems incurred by the large influx of x-rays at the start of each round led NIOSH to discontinue this administration procedure. Instead, to smooth out processing, mine plans were filed on a continuous basis after 1981. In this report, for convenience, the period between August 1981 and December 1986 will be referred to as round 4, although it is not a round in the same sense as were the previous three. Details pertinent to each round are given in Table 1.

Film Readings

The results presented here are confined to data based on the first and second x-ray readings. (Results are not given for the final determinations because the procedures by which they were obtained differed between the rounds.) At least two readings were made on the films from rounds 1, 3, and 4. Because of logistical problems at round 2, the decision was made not obtain a second reading if the first reading had been made by a B reader (10% of all films). In rounds 3 and 4, the second reading was always made by a B reader, but both A and B readers performed the first reading. Three sets of x-ray standards were used over the four rounds: the Union Internationale Contre le Cancer (UICC)/Cincinnati standards of 1968[7] (round 1), the 1971 International Labour Office (ILO) standards[8] (rounds 2 and 3), and the 1980 ILO standards[9] (round 4). Results are presented on prevalences of small rounded and small combined opacity profusions of progressive massive fibrosis. (For details pertaining to the definition of CWP and the derivation of small rounded opacity scores, see the report in this Journal on epidemiological findings.)[10]

In the following analysis, there are fewer readings than participants owing to the unavailability of data for films said to be unreadable. In some cases, multiple films for the same person were available within the same round; in these cases, only readings of the last film are included.

Results

The following analysis provides information on tenure-specific prevalences among the participants, and on summary statistics computed to facilitate the comparison of prevalences over the different rounds.

Tenure Distribution Data

Table 2 provides detailed information on the tenure distribution of participants at each round. Occupational history information was obtained from the miners by

interview at the time of examination. Note that the data from each round are probably weighted toward younger miners; this is owing to the dual mandatory/voluntary nature of the program. New miners, who must have initial and 3-year x-rays, tend to be younger. The experienced miners, who are generally older and not required to obtain regular x-rays, have participated less in the program. As a result, the tenure distributions of CWXSP participants might be expected to be different from those that would have resulted from a census of the work force held at the same time.

Furthermore, selective participation may have occurred, with the healthier miners having been more likely to choose to attend. Lastly, the age structure of the mining work force has been changing dramatically since 1969. In particular, there was a large reduction in the average age between 1969 and 1979 as older men retired from the generally elderly work force, and young men were hired to fill the deficit and to man mines newly opened or expanded in response to the energy crisis of the mid-1970s.

The table also includes a projected tenure distribution for the 1986 work force for underground and surface workers at underground coal mines; this distribution was derived from a mail survey of coal mine operators.[11]

TABLE 1—Background Information on Each Round

	Round 1	Round 2	Round 3	Round 4
Dates	1970–1973	1973–1978	1978–1981	1981–1986
X-ray classification system	1968[7]	1971[8]	1971[8]	1980[9]
Number of participants	77 756	122 625	63 519	40 634
Average total population in underground coal mining employment	103 000	128 000	144 000	110 000

Note. Work force estimates are from MSHA reports, such as that for 1988.[10]

TABLE 2—Tenure Distributions for Miners Who Attended the Four Completed Rounds of the Coal Workers' X-ray Surveillance Program and for Those Included in the 1986 Survey of Coal Miners[11]

Tenure in Mining Group	% Round 1 Miners (n = 71 404)	% Round 2 Miners (n = 107 031)	% Round 3 Miners (n = 58 282)	% Round 4 Miners (n = 34 953)	% 1986 Survey Miners[a] (n = 78 741)
0[b]	22.1	46.7	24.9	17.1	2.0
1	7.4	7.2	6.4	3.5	2.0
2–4	11.9	14.8	20.7	16.2	8.9
5–9	9.4	10.8	24.3	28.9	29.2
10–14	6.2	4.4	9.1	19.8	35.0
15–19	6.6	2.9	3.7	6.6	14.1
20–24	10.2	2.8	2.6	2.6	3.6
25–29	8.8	4.0	2.2	1.6	1.6
30+	17.8	6.5	6.1	3.7	3.7

[a]The tenure figures for the 1986 survey were computed from data supplied by the Bureau of Mines. Because the original data were rounded up to the nearest integer, 1 year was subtracted from the tenure to make the data consistent with the radiological data (in which the figures were rounded down).
[b]The zero tenure group includes all those with less than 1 year.

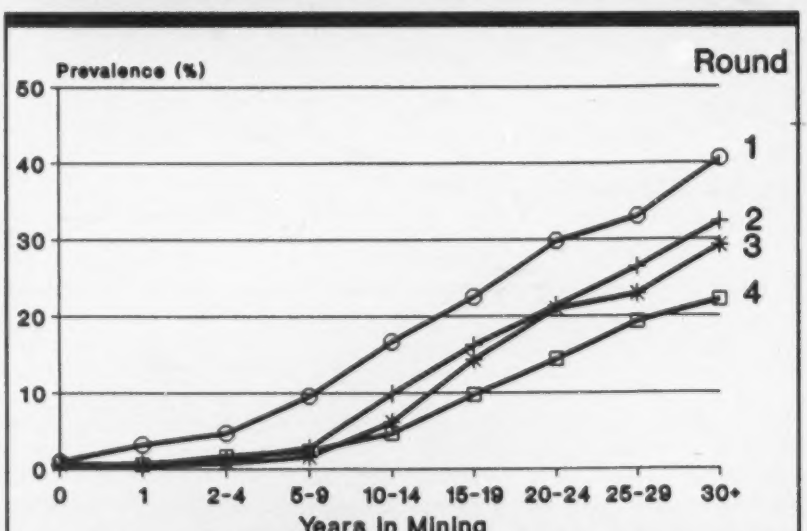


FIGURE 2—Tenure-specific prevalences of category 1 or greater small rounded opacities by round of the Coal Workers' X-ray Surveillance Program, based on the first readings only.

Tenure-Specific Prevalences: Small Rounded Opacities

Figures 2 and 3 show tenure-specific prevalence rates of category 1 or greater small rounded opacities for the first and second readers, respectively. Comparison of these tenure-specific rates across rounds reveals several interesting points. The first readings (Figure 2) show a consistent decline in prevalence over all four rounds in virtually every tenure group. In comparison, the second readings (Figure 3) reveal higher prevalences in round 2 for longer-tenured miners; otherwise, the trends are similar. Comparison of the two figures reveals that, at all four rounds, the second readers tended to report fewer cases of abnormality than the first readers, an effect that is most obvious in round 1.

Figures 4 and 5 provide corresponding statistics for category 2 or greater small rounded opacities. Both the first and second readings show a consistent reduction in prevalence over time; however, the general prevalence levels are higher for the first readings.

Summary Statistics for Small Rounded Opacities

The data given in Figures 2 through 5 were used to derive certain summary statistics. First, summation of the products of tenure-specific prevalence shown in those figures with the tenure distribution proportions from Table 2 for each round gave rise to crude prevalence rates for all participants by round (see Table 3). How-

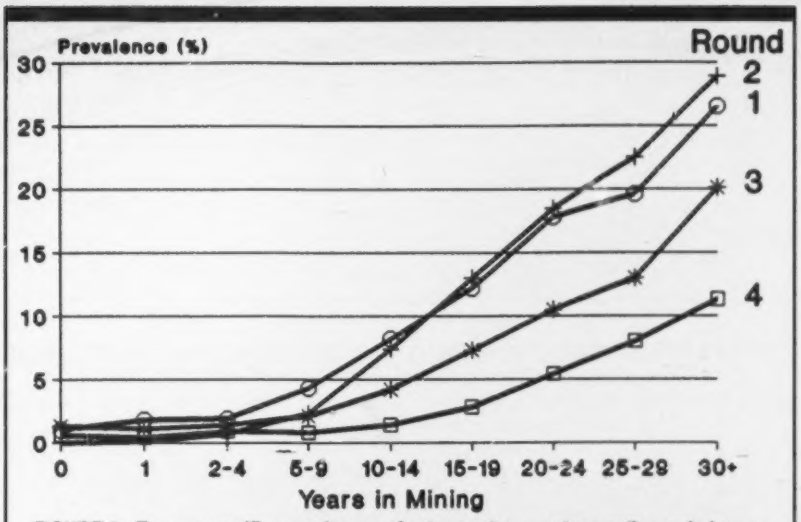


FIGURE 3—Tenure-specific prevalences of category 1 or greater small rounded opacities by round of the Coal Workers' X-ray Surveillance Program, based on the second readings only.

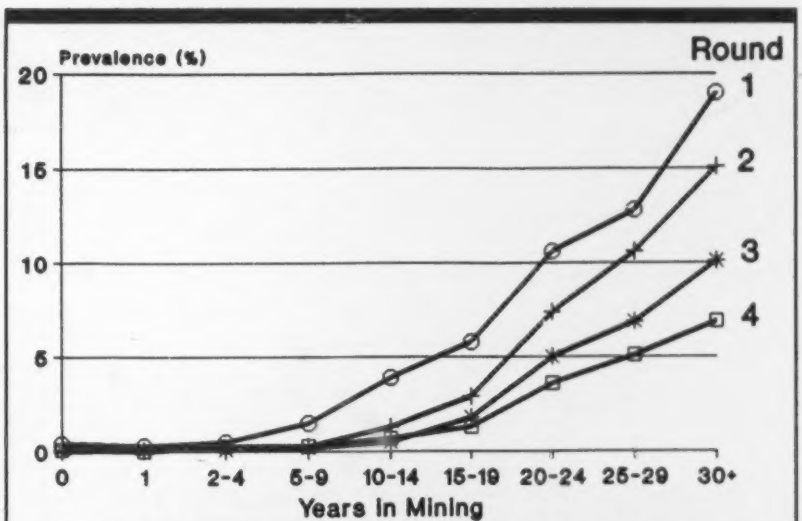


FIGURE 4—Tenure-specific prevalences of category 2 or greater small rounded opacities by round of the Coal Workers' X-ray Surveillance Program, based on the first readings only.

ever, these rates are naturally heavily weighted by the zero tenure group and, to some extent, by the 3-year x-rays. Thus, to eliminate the spurious and misleading weighting caused by the mandatory preemployment and 3-year x-rays, the above approach was modified by omitting the 0- to 4-year tenure group from the summation and adjusting to account for the fact that the tenure proportions no longer add to 100%. Not surprisingly, these statistics are all larger than the crude rates (see Table 3). They tend to reflect the influence of changes in both employment patterns and dust control, while removing the effect of the mandatory films.

The question of whether dust control brought about a reduction in prevalence after 1969 was tackled by standardizing the rates for each round to a constant tenure distribution. A summary statistic was formed by multiplying tenure-specific prevalences from Figures 2 through 5 for each round by the proportions of miners in the tenure groups given by the Bureau of Mines survey.[11] Note that these statistics were generated for interround comparison and should not be regarded as estimates of prevalence at any particular round of the CWXSP, except perhaps the fourth. Other tenure distributions could obviously have been used for standardization.

Overall, the summary statistics reveal that the first readers consistently reported more abnormality than did the second readers. The summary rates also show that, once the weighting effect of the new miners is removed (e.g., by restricting the calculations to those miners with 5 or more years of tenure), the prevalences show a much more gradual and uniform drop from round 1 to round 4. Hence, large temporal changes in tenure distribution were clearly responsible for large variations in the crude summary prevalence statistics. Lastly, standardization to a common tenure distribution leads to the smallest drop in prevalence from round 1 to round 4 of the three summary statistics, although clear downward trends over the rounds are still evident.

Small Combined Opacities

Although most research into the radiographic appearances of CWP has concentrated on rounded opacities, there have been some recent indications that irregular opacities are also associated with dust exposure.[12] For completeness, the prevalences of both rounded and irregular opacities (i.e., combined opacities) have been tabulated, with the resulting adjusted rates shown in Table 4. As expected, the rates are all higher but otherwise follow the same trends seen in Table 3.

Prevalence of Large Opacities

Standardization for tenure was not undertaken on the data for large opacities because there were too few cases. However, tenure-specific prevalences of progressive massive fibrosis by round and reader are given in Figures 6 and 7. The data exhibit similar tendencies to those seen for small opacities—namely, a rise in prevalence with tenure and a trend toward lower prevalences of disease at the later rounds.

Participation

Participation is obviously an important issue with these data: if it is not 100%, self-selection and, hence, bias might be suspected. Unfortunately, assessing bias in the CWXSP to any detailed extent is very difficult owing to a paucity of data by age, region, etc., on the baseline population of miners eligible for the program at any time.

Some very rough rates of participation pertinent to miners eligible for the vol-

untary aspect of the program were derived. This was done by dividing the numbers of examined miners with 1 or more years of tenure at each round by average work force estimates applicable to each round obtained from annual Mine Safety and Health Administration reports on injuries, such as that for 1986.[13] The results indicate that participation of miners with 1 or more years of tenure at rounds 1 through 4 was 50%, 44%, 32%, and 30%, respectively.

Some indication of tenure-specific participation can be gained for round 4 of the CWXSP by using some Bureau of Mines survey data[11] (shown in Table 2 of this report). Tenure-specific participation rates based on these data were 44%, 25%, 21%, 32%, 44%, and 44% for the 5–9, 10–14, 15–19, 20–24, 25–29, and 30+ groups, respectively. It is interesting that participation appears to improve among miners with longer tenure in mining.

Comparison with Epidemiological Data

Undertaken concurrently with the coal miner surveillance program was a large-scale, nationwide epidemiological study of underground coal miners. Results from that research are described in a companion report.[10] Summary rates of small opacity profusion standardized to the Bureau of Mines survey tenure distribution were derived (see Tables 3 and 4); these reveal trends similar to those noted here for the surveillance data, although the reported prevalences appear to be generally lower.

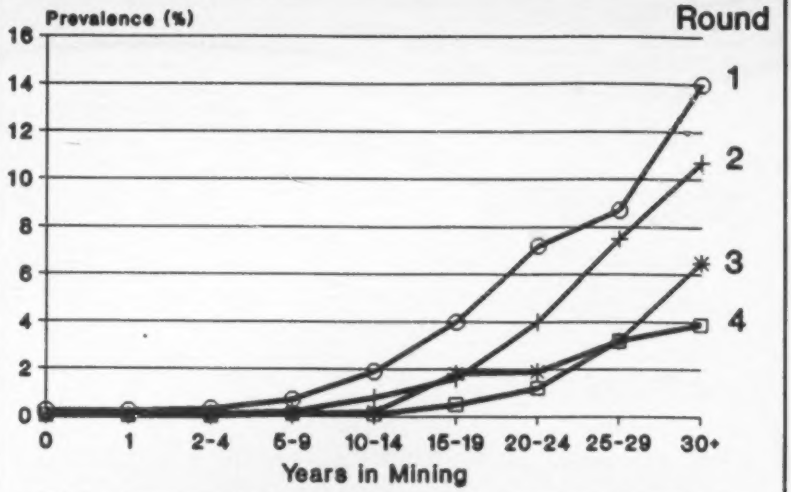


FIGURE 5—Tenure-specific prevalences of category 2 or greater small rounded opacities by round of the Coal Workers' X-ray Surveillance Program, based on the second readings only.

TABLE 3—Adjusted Summary Prevalence Estimates for Small Rounded Opacities by Round of the Coal Workers' X-ray Surveillance Program, Separately by First and Second Readers, and by Data from a Large Concurrent Epidemiological Study[10]

		Adjusted Summary Prevalences, %			
	Category[a]	Round 1	Round 2	Round 3	Round 4
First readers					
All participants[b]	1+	17.7	5.4	4.7	4.2
	2+	5.5	1.5	0.9	0.7
Tenure >4 years[c]	1+	28.2	15.8	9.0	6.0
	2+	9.2	4.7	1.7	1.1
Common tenure distribution[d]	1+	15.4	9.1	6.9	5.5
	2+	3.4	1.5	0.9	0.9
Second readers					
All participants[b]	1+	10.5	4.4	3.6	1.8
	2+	3.7	0.9	0.5	0.2
Tenure >4 years[c]	1+	16.9	13.5	6.1	2.2
	2+	6.2	3.0	1.1	0.4
Common tenure distribution[d]	1+	8.0	7.3	4.6	2.0
	2+	2.0	1.0	0.5	0.3
Epidemiological data					
Common tenure distribution[d]	1+	4.5	4.1	1.8	1.3
	2+	1.2	1.0	0.4	0.1

[a]1+ means category 1 or greater; 2+ means category 2 or greater.
[b]All participants = summary rates based on all mandatory and voluntary x-rays.
[c]Tenure >4 years = summary rates based on all miners with more than 4 years tenure in mining.
[d]Common tenure distribution = summary rates standardized to data in right-most column of Table 2.

Discussion

The information reported here reveals a clear downward trend in CWP prevalence among program participants over time. The decline parallels that observed over a series of medical surveys undertaken as part of a long-term epidemiological study of US underground coal miners[10] (see Tables 2 and 3 of this report). Both sets of data suggest that prevalences of category 1 or greater small opacities at round 4 (1981 to 1986) were about one-third of their levels at round 1 (1970 to 1973). Reductions in reported prevalence of category 2 or greater small opacities were somewhat larger.

These reductions in prevalence coincide with lower dust exposure limits mandated by the 1969 Federal Coal Mine Health and Safety Act.[1] Under the regulations promulgated by that legislation, permissible airborne respirable dust concentrations in underground coal mines are currently limited to 2 mg/m^3. This contrasts to levels of 6 mg/m^3 or more existing before 1969.[14] Documented trends of dust concentrations in underground mines indicate a substantial decline over the 10 years from 1968,[15] although there is some evidence that the reductions have not been as great as reported.[16,17]

Because the 1986 underground coal mining work force was fairly young, use of this distribution for standardization leads to rates reflecting prevalence among the less experienced miners. This is appropriate if preliminary evidence on the efficacy of the current federal dust limit is being sought, for it is only those with limited tenure who have, to date, been working under the lower dust

TABLE 4—Adjusted Summary Prevalence Estimates for Small Combined Opacities by Round of the Coal Workers' X-ray Surveillance Program, Separately by First and Second Readers, and by Data from a Large Concurrent Epidemiological Study[10]

	Category[a]	Adjusted Summary Prevalences, %			
		Round 1	Round 2	Round 3	Round 4
First readers					
All participants[b]	1+	22.4	7.1	6.0	5.5
	2+	6.5	1.8	1.1	0.8
Tenure >4 years[c]	1+	35.0	20.3	11.4	7.8
	2+	10.8	5.7	2.2	1.2
Common tenure distribution[d]	1+	19.5	11.7	8.7	7.2
	2+	4.0	2.0	1.2	1.0
Second readers					
All participants[b]	1+	13.8	5.9	5.7	3.0
	2+	4.5	1.2	0.6	0.3
Tenure >4 years[c]	1+	22.1	18.2	9.2	4.0
	2+	7.5	3.8	1.3	0.5
Common tenure distribution[d]	1+	10.7	9.9	7.3	3.6
	2+	2.4	1.2	0.7	0.4
Epidemiological data					
Common tenure distribution[d]	1+	6.6	5.1	3.6	2.3
	2+	1.5	1.2	0.5	0.3

[a]1+ means category 1 or greater; 2+ means category 2 or greater.
[b]All participants = summary rates based on all mandatory and voluntary x-rays.
[c]Tenure >4 years = summary rates based on all miners with more than 4 years tenure in mining.
[d]Common tenure distribution = summary rates standardized to data in right-most column of Table 2.

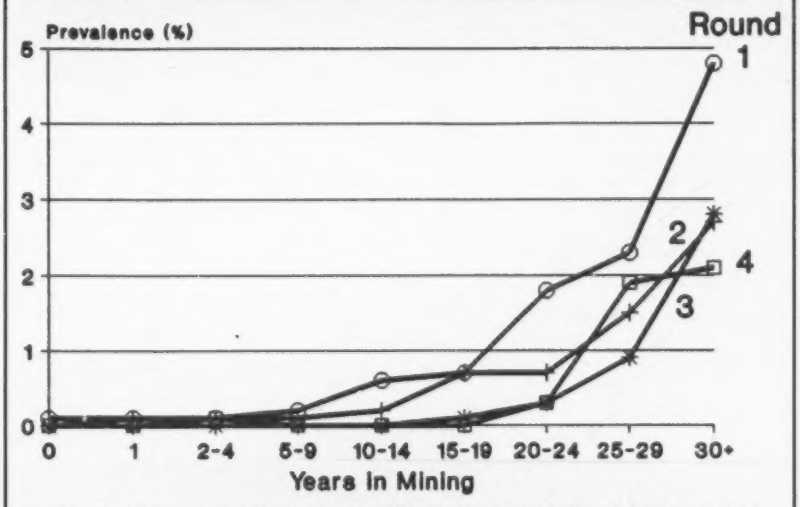


FIGURE 6—Tenure-specific prevalences of large opacities by round of the Coal Workers' X-ray Surveillance Program, based on the first readings only.

limits mandated by the 1969 law. However, this distribution is inadequate for estimation and comparison of prevalence levels and changes in older miners. Simple averages of the tenure-specific rates for miners with 20 or more years of work in mining indicate drops for category 1 or greater small rounded opacities from 34.5% to 23.3% for the first readers and from 12.1% to 5.7% for the second readers. Corresponding figures for category 2 or greater small rounded opacities were from 24.4% to 8.2% for the first readers and from 6.2% to 2.2% for the second readers. As might be expected, the reductions in prevalence (as measured by the ratios of the prevalences) are not so marked as those obtained using the 1986 work force distribution (Table 3, common tenure distribution). This may be because many of the older miners in round 4 had substantial lung dust burdens that were gained from exposures received before 1969.

In general, the prevalences estimated from the surveillance data are greater than those from the epidemiological study. It is not clear why this occurred. Possible reasons for the difference may lie with x-ray reader effects or with systematic differences in prevalence associated with the types and sizes of the mining operations feeding into the two programs. The epidemiological study has always involved large operations, which may have better dust-control procedures and thus lower levels of CWP than the smaller mines. Smaller mines form a large part of the input to the CWXSP.

Prior reports on the CWXSP have generally concentrated on certain restricted aspects of the data.[18-21] The findings from those papers are essentially the same as those presented here and are not discussed further.

There are obviously many problems in interpreting data from successive cross-sectional surveys. These include exodus of workers with ill-health, self-selection bias, changes in the x-ray classification standards, and interreader variability. There is also the possibility that readers' perceptions of CWP have been changing over time. Some reading exercises involving large amounts of film are planned. These should provide more reliable information on trends in prevalence of CWP in underground coal miners.

Longitudinal study approaches can control for many of the factors that interfere with interpreting cross-sectional data. In such studies, a specific cohort of workers is followed over time, with the x-rays being read at one time in a controlled reading environment. Althouse et al.[22] conducted one such study using films from the CWXSP and looking at x-ray changes over 10 years in new miners. Of 1825 miners studied, a net total of 15 developed category 1 opacities from category 0; no category 2 or greater readings were reported. While these reported incidence levels are encouraging, they should be regarded as only preliminary; further studies are needed owing to the relatively short observation period and to the fact that CWP is often not manifested until after more than 10 years of exposure.

In conclusion, despite methodological limitations, the results presented here demonstrate a temporal reduction in prevalence of CWP that is consistent with results from concurrent epidemiological investigations and coincides with reductions

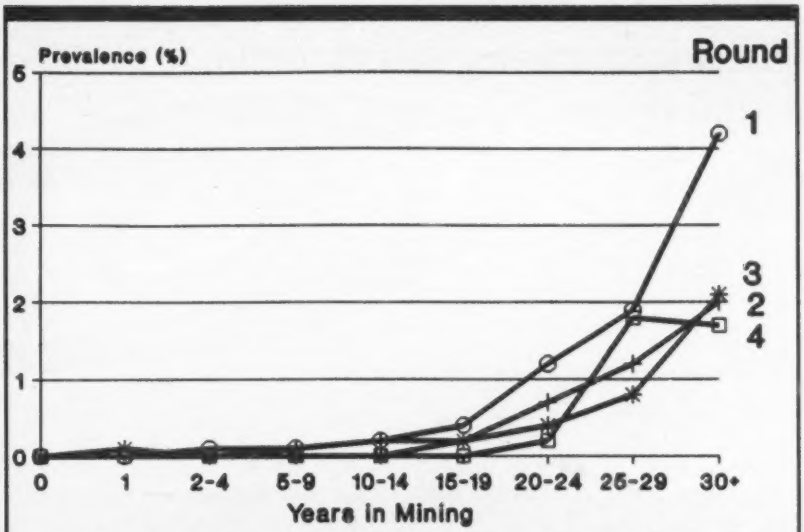


FIGURE 7—Tenure-specific prevalences of large opacities by round of the Coal Workers' X-ray Surveillance Program, based on the second readings only.

in dust levels mandated after 1969. Continual vigilance in dust control is necessary to ensure that these gains are maintained and, it is hoped, advanced. Emphasis on dust control is especially important with regard to the increasing use of mining techniques that tend to cause greater dust generation, such as longwall methods.[23] In addition, the implications of coal mine dust exposure for lung diseases other than pneumoconiosis must be fully determined. □

Acknowledgments

Thanks are due to Barbara Bonnett for help in preparation of the figures and in other programming, and to the many who have helped in preparing and reviewing the manuscript.

References

1. Coal Mine Health and Safety Act. US Pub L No. 91–173, Stat 2917.
2. 42 CFR PT 37. Specifications for medical examinations of underground coal miners. 35 Fed Reg 13 206-13 208 (1970).
3. 42 CFR PT 37. Specifications for medical examinations of underground coal miners. 38 Fed Reg 20 076-20 081 (1973).
4. 42 CFR PT 37. Specifications for medical examinations of underground coal miners. 43 Fed Reg 33 713-33 720 (1978).
5. 42 CFR PT 37. Specifications for medical examinations of underground coal miners. 44 Fed Reg 23 084-23 085 (1979).
6. Morgan RH. Proficiency examination of physicians for classifying pneumoconiosis chest films. *Am J Radiol.* 1979;132:803–808.
7. Bohlig H, Bristol LJ, Cartier PH. UICC/Cincinnati classification of the radiographic appearances of categories of simple pneumoconiosis: a cooperative study by the UICC committee. *Chest.* 1970;58:57–67.
8. International Labour Office. ILO U/C international classification of radiographs of pneumoconioses 1971. Occupational Safety and Health, Series no. 22 (Rev). Geneva, Switzerland: International Labour Office; 1972.
9. International Labour Office. *International Classification of Radiographs of Pneumoconiosis.* Rev ed. Occupational Safety and Health, Series no. 22 (Rev 80). Geneva, Switzerland: International Labour Office; 1980.
10. Attfield MD, Castellan RM. Epidemiological data on US coal miners' pneumoconiosis, 1960 to 1988. *Am J Public Health.* 1992;82:964–970.
11. Butani SJ, Bartholomew AM. *Characterization of the 1986 Coal Mining Work Force.* Washington, DC: US Bureau of Mines; 1988. Bureau of Mines Information Circular 9192.
12. Collins HPR, Dick JA, Bennett JG, et al. Irregularly shaped small shadows on chest radiographs, dust exposure, and lung function in coalworkers' pneumoconiosis. *Br J Ind Med.* 1988;45:43–55.
13. Mine Safety and Health Administration. *Injury Experience in Coal Mining, 1986.* Washington DC: US Govt Printing Office; 1987. Information Report 1157.
14. Jacobson M. Respirable dust in bituminous coal mines in the US. In: Walton WH, ed: *Inhaled Particles III.* Old Woking, England: Unwin Brothers; 1971:745–756.
15. Parobeck PS, Jankowski RA. Assessment of the respirable dust levels in the nation's underground and surface coal mine operations. *Am Ind Hyg Assoc J.* 1979;40:910–915.
16. Boden LI, Gold M. The accuracy of self-reported regulatory data: the case of coal mine dust. *Am J Ind Med.* 1984;6:427–440.
17. Mine Safety and Health Administration. Peabody Coal Company pleads guilty to dust sample tampering. Washington DC: Mine Safety and Health Administration; January 17, 1991. Press release USDL 91–23.
18. Felson B, Morgan WKC, Bristol LJ, et al. Observations on the results of multiple readings of chest films in coal miners' pneumoconiosis. *Radiology.* 1973;109:19–23.
19. Merchant JA. Coal workers' pneumoconiosis surveillance in the United States. In: *Proceedings of the Fifth International Conference on Pneumoconiosis.* Caracas, Venezuela: 1978.
20. Millar JD. *Statement before the House Subcommittee on Oversight, Committee of Ways and Means,* July 27, 1981.
21. Althouse RB. Ten years experience with the national coal workers health surveillance program. *MMWR CDC Surveill Summ.* 1985;34/1SS:33–37SS.
22. Althouse RB, Attfield MD, Kellie S. Use of data from the x-ray surveillance program for underground coal workers to evaluate the efficacy of the dust standard, presented at Conference on Medical Screening and Biological Monitoring for the Effects of Exposure in the Workplace, Cincinnati, July 1984. *J Occup Med.* 1986; 28:741–745.
23. Watts WF, Niewiadomski GE. Respirable dust trends in coal mines with longwall or continuous miner sections. In: *Proceedings of the Seventh International Pneumoconiosis Conference, Pittsburgh, USA, 1988.* Washington, DC: US Department of Health and Human Services; 1990:94–99. DHHS (NIOSH) publication 90–108, part I.

British Data on Coal Miners' Pneumoconiosis and Relevance to US Conditions

Michael D. Attfield, PhD

ABSTRACT

Objectives. The current primary federal dust standard for US underground coal miners of 2 mg/m^3 respirable dust is based on British epidemiological information on exposure-response derived in 1969. Since then, much new information has become available. This paper reviews and compares the available information as it relates to the US mining situation.

Methods. Recent exposure–response information on pneumoconiosis and dust exposure derived by British researchers was employed to estimate working-life risks of pneumoconiosis for miners exposed to 2 mg/m^3.

Results. It is estimated that close to 9% of underground coal miners who work for 40 years in a 2 mg/m^3 environment would develop pneumoconiosis (category 1 or greater). Progressive massive fibrosis would develop in 0.7%.

Conclusions. There are unresolved questions relating to the validity of extrapolating findings on British mines and miners to the US and also in predicting disease levels at the low end of the dust exposure spectrum. Given the data available, current information suggests miners who are employed for a working lifetime at the current federal dust limit of 2 mg/m^3 are still at risk of developing pneumoconiosis. (*Am J Public Health.* 1992;82:978–983)

Introduction

It was not until 1969 that a nationwide limit was set on the level of dust to which US underground coal miners could be exposed. In that year, the Federal Coal Mine Health and Safety Act was passed.[1] This mandated that a dust standard of 3 mg/m^3 be imposed, with a reduction to 2 mg/m^3 to come into force in December 1972. (There is currently a requirement that dust levels be reduced further if the mine dust contains more than 5% quartz.) The intention was to prevent premature death and disability associated with coal workers' pneumoconiosis (CWP).

Owing to the lack of adequate epidemiological information on US underground coal miners at the time, those responsible for the law relied heavily on findings from overseas studies, primarily some interim results from Britain. Over the 23 years that have elapsed since the passage of the act, the British researchers have pursued their analyses further and developed more refined analyses. The purpose of this paper is to review both the initial and later results from Britain, and to discuss their applicability to the US underground coal mining situation.

After some preliminary background information, the paper outlines the scientific basis of the current federal 2 mg/m^3 standard. It then reviews the epidemiological findings that have been published since 1969 on British miners. Finally, it discusses some issues that involve the application of these results to US coal miners.

Note that this report is concerned only with CWP as determined through examination of chest x-rays; it does not deal with questions relating to the utility of the current dust standard in preventing other diseases, such as bronchitis and emphysema. Furthermore, no attempt has been made here to assess the efficacy of current disease-prevention measures through examination of CWP prevalence rates in US coal miners observed since 1969. Prevalence information is available in two reports: one on data from a national x-ray screening program known as the Coal Workers' X-ray Surveillance Program,[2] and one on results from a large continuing epidemiological study called the National Study of Coal Workers' Pneumoconiosis.[3]

Background

Coal mining existed in the United States for many years before it was accepted that many coal miners were suffering from occupational lung disease; and although silicosis was known to occur in some miners, the inhalation of coal dust with a low silica component was often thought to be harmless or even beneficial.[4] However, the accumulation of medical evidence to the contrary eventually led in 1969 to the enactment of federal regulations for the control of coal mine dust.

CWP is divided into two disease states: simple CWP and complicated pneumoconiosis (also called progressive massive fibrosis, or PMF).[5] In living coal miners, simple and complicated CWP are both almost always detected through examination of chest x-rays. The severity of simple CWP is measured in terms of the profusion of opacities seen on the x-ray

Michael D. Attfield is with the Division of Respiratory Disease Studies at the National Institute for Occupational Safety and Health in Morgantown, WV.

Requests for reprints should be sent to Michael D. Attfield, PhD, MS 234, ALOSH, 944 Chestnut Ridge Road, Morgantown, WV 26505

This paper was submitted to the Journal September 14, 1990, and accepted with revisions March 19, 1992.

and is recorded according to major categories (0, 1, 2, or 3), with each category being divided into three subcategories depending on the adjacent major category (e.g., major category 2 is divided into 2/1, 2/2, and 2/3). In this report, the emphasis is on rounded opacities (types p, q, and r), although recent evidence also suggests a relationship between irregular-shaped opacities (types s, t, and u) and coal dust exposure.[6] PMF is defined by the presence of one or more pneumoconiotic opacities of at least 1 cm in diameter. Miners who have simple CWP are at greatly increased risk of developing PMF, and this risk increases with category of simple CWP.[7]

In addition to the imposition of a dust limit, the federal act provided for a complementary radiographic screening program to identify miners who, for reasons of overexposure or individual susceptibility to dust, showed signs of CWP. Such miners are given the option of working in a low-dust (1 mg/m^3) environment while retaining their pay and benefits.

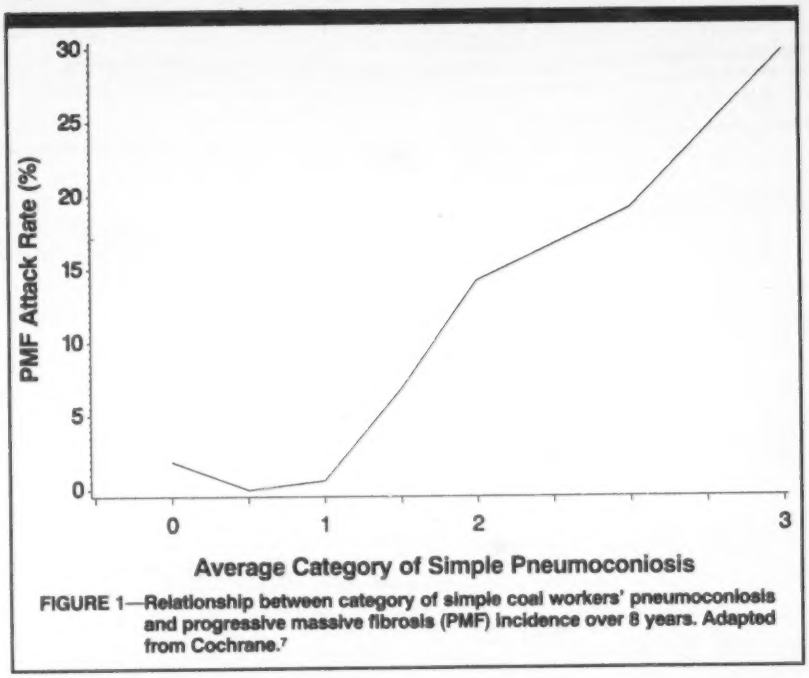


FIGURE 1—Relationship between category of simple coal workers' pneumoconiosis and progressive massive fibrosis (PMF) incidence over 8 years. Adapted from Cochrane.[7]

British Studies of CWP and Dust Exposure

British studies of CWP and other lung diseases associated with coal mining began in 1952. The principal program of investigation was the Pneumoconiosis Field Research (PFR).[8] In the PFR, coal miners working at mines located in all of Great Britain's major coalfields were medically examined at roughly 5-year intervals until the 1980s. Moreover, industrial hygienists were stationed at all the mines, their purpose being to assess the dust exposures of all miners prospectively as the study progressed. A systematic and intensive program of dust sampling was undertaken on all occupational groups, and the resulting mean dust concentrations were eventually linked with job tenure information gathered on each miner. Finally, respirable dust exposures that were developed for each miner were correlated with indices of lung disease in various studies, some of which are the subject of the following discussion.

The PFR must rank as one of the best conceived and managed epidemiological studies. Its features are a large number of subjects studied (initially 35 000 in 25 mines); continuous efforts to standardize data collection techniques and to minimize all sources of variability; a prospectively based program of dust exposure assessment, which was undertaken independently of any compliance sampling and was operated according to a scientifically designed plan of dust sampling; the collection of detailed time records of work performed by each miner; and excellent participation by miners in the study.

Basis for the Current Federal Dust Standard

In 1969, when the Federal Coal Mine Health and Safety Act was enacted in the United States, prevention of the disabling disease PMF was naturally one of its principal objectives. However, no information existed at that time that directly linked the degree of dust exposure to the risk of developing the disease. Nevertheless, two other important pieces of knowledge of relevance to this problem were available.

First, it was known that the risk of developing PMF was considerably greater among miners who had advanced simple CWP, such as category 2 or 3 (see Figure 1).[7] The abrupt rise in PMF risk associated with the higher categories of simple CWP suggested that PMF could be largely eliminated if progression to those categories could be prevented. Hence, it appeared desirable to limit the dust to a level where progression of miners to category 2 or greater simple CWP did not occur.

Second, based on data from the PFR, a curve had recently been developed, which related the risk of developing simple CWP category 2 or greater over 35 years to average dust concentration (Figure 2).[9] This curve was not derived from observation of incidence over that period but was actually extrapolated from results of a 10-year study, the Interim Standards Study (ISS).[10] From the 35-year curve, it was possible to identify the level of dust that would lead to any given degree of incidence of category 2 simple CWP or greater. Examination of the curve indicated that an exposure to 2 mg/m^3 coal mine dust would lead to zero incidence. Hence, it was argued, if exposures were kept to 2 mg/m^3 or below, miners would not progress to category 2 and thus would never get to be at high risk of developing PMF. It was on this basis that the federal standard was set at 2 mg/m^3. The Coal Workers' X-ray Surveillance Program was intended as a backup prevention strategy to identify those workers who might develop CWP despite dust control, and to enable such workers to work from then on in a low-dust environment.

Recent Information from the PFR on Exposure-Response for CWP

Although the strategy of preventing PMF through prevention of severe simple CWP was a reasonable course of action in 1969, given the lack of information that existed then on a direct relationship between PMF and dust exposure, it has a basic weakness. This was revealed in a

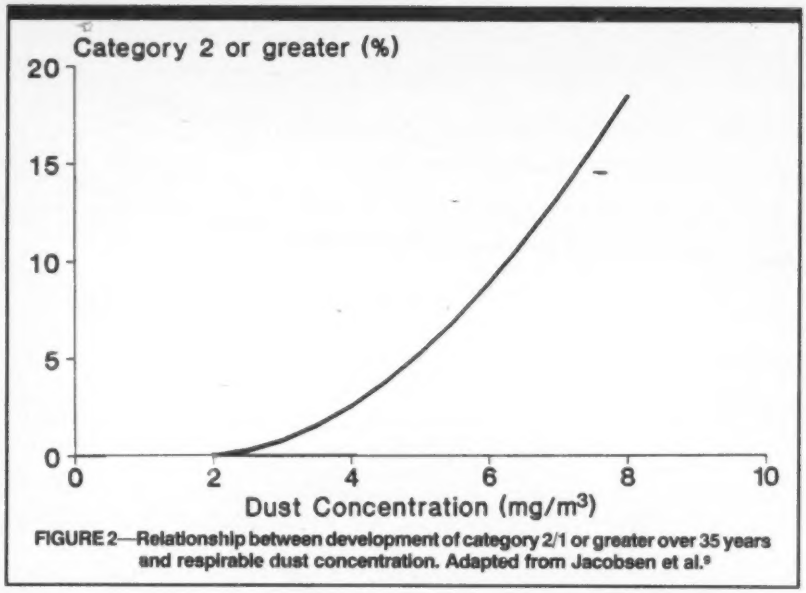


FIGURE 2—Relationship between development of category 2/1 or greater over 35 years and respirable dust concentration. Adapted from Jacobsen et al.[9]

series of articles that examined factors connected with the incidence of PMF.[11–14] The reports indicated that, contrary to Cochrane's findings,[7] the risk of PMF appeared somewhat lower for miners with categories 2 and 3 simple CWP but was rather higher for those with category 1. Because many more miners get category 1 CWP compared with those who get category 2 or 3, this implies that prevention of category 2 or greater would not eliminate the risk of getting PMF. In this respect, Shennan et al.[12] found that half the cases of PMF occurring in British miners over a 4- to 5-year period in the late 1970s were seen in miners who had category 0 or 1 at the initial examination. These findings suggested the need for an alternative approach to the prevention of PMF.

This new strategy was provided in the form of a curve linking PMF incidence directly to levels of dust exposure reported in 1985.[15] This curve was derived from data collected from 1959 to 1977 on over 30 000 British coal miners. Five-year estimates of the risk of PMF attack were obtained based on age and dust exposure; these were then compounded into 35-year predictions for various dust levels. The resulting curve indicated that an exposure to 2 mg/m^3 dust was associated with about a 1% risk of PMF. That is, 1 out of every 100 miners who worked for 35 years in an atmosphere containing 2 mg/m^3 of coal mine dust would be expected to develop PMF.

Although the above results suggest that coal miners who work for 35 years in a 2 mg/m^3 environment may be at risk of PMF, it should be noted that the analysis did not take into account the effect of secondary preventive measures—namely, x-ray surveillance and job transfer. In theory, people who exhibit early signs of CWP and then transfer to a low-dust environment should reduce their risk of progressing to a higher category of simple CWP and thus lower their chance of contracting PMF. To explore the effect of both primary and secondary measures on disease incidence in US coal miners, PFR researchers prepared a special analysis.[16] Their object was to determine what the incidence of PMF would be if a miner were to start work at age 18 and work until age 58 (a typical retirement age in US coal miners) in an atmosphere containing 2 mg/m^3 of coal mine dust, with transfer to a 1 mg/m^3 environment for those miners who were found by x-ray surveillance to have CWP category 1 or greater.

Latest Predictions of CWP Risk

Full details of the data and methods used to derive these predictions are provided by Hurley and Maclaren.[16] In brief, the results are based on virtually the same set of data used by Hurley and Jacobsen[15] and also described elsewhere,[13,14] but they included coal rank (percent carbon) as a predictor variable and used a refined method to extrapolate for lifetime risk. Results were obtained from observation of 52 264 approximate 5-year man-intervals of risk for 30 500 miners having reliable dust exposure records and x-ray readings over the period of 1953 to 1977.

Each man-interval of risk gave rise to four responses based on radiographic status at the end of the interval (category 1, 2, or 3, and PMF, category 0 being obtained by subtraction). These responses were modeled using logistic methods against initial category, age, dust exposure up to the initial survey, and coal rank (percent carbon). Support for the inclusion of this variable as well as the others can be found in earlier analyses of these data.[13] Predicted probabilities of x-ray change for 5-year age groups based on these models were then obtained by substituting various levels of dust exposure and percent carbon. (For example, 83% is an average value for the common US type "A" high volatile bituminous coal.) These 5-year probabilities could then be compounded into predictions for various ages up to age 58, assuming that every miner started in category 0 at age 18 (see Hurley and Maclaren[16] for the exact details). The effect of secondary preventive measures was allowed for by assuming that all those who were found to have category 1/0 or greater immediately worked from then on in a 1 mg/m^3 environment, as allowed for by the 1969 act.

Predictions of PMF and simple CWP levels for US miners obtained using these methods for a dust concentration of 2 mg/m^3 with 83% carbon are shown in Table 1 by age and category of CWP. They indicate that 0.7% of miners who worked their full working lives in coal mining at the present maximum of 2 mg/m^3 dust (1 mg/m^3 after contracting CWP) would have PMF at age 58, while close to 2.4% would have category 2 or higher (including PMF). Figure 3 shows two curves, one for PMF and one indicating the percentage of miners predicted to have category 2 or greater simple CWP at age 58 after 40 years of mining at dust levels ranging from 1 to 8 mg/m^3 at 83% carbon. (I computed the predicted values in Table 1 and Figure 3 from the results of Hurley and Maclaren, because they presented predictions only for 2 mg/m^3.)

The above predictions apply only to type "A" high volatile bituminous coal, a type of coal that is mined generally west of the eastern divide of the United States. They do not apply to the anthracite region in eastern Pennsylvania and to certain other areas, such as central Pennsylvania and counties in the southeast of West Virginia, where low volatile bituminous coal is mined. In both of those regions, the coal rank is higher, attaining levels of 89% or greater; and the risk of PMF according to the model is correspondingly higher, being estimated at about 1.8% for miners with 40 years of tenure at 2 mg/m^3 dust in

low volatile coals. The risk of category 2 or greater is estimated to be close to 5% in such miners. The predicted risks for anthracite workers are higher still.

Differences between Predictions

The predictions for category 2 or greater simple CWP that were derived from the various exposure-response curves show broad agreement. For 35 years' exposure, the predictions range from 0% for the ISS analysis to between 1.24% and 1.61% for the results derived by Hurley and Maclaren[16] and summarized in Table 1 (interpolated at age 53 for this comparison), depending on whether PMF is included (the ISS analysis did not take PMF into account). For the two studies involving PMF, the agreement is closer, the estimated prevalences at 2 mg/m^3 being about 1%, from Hurley and Jacobsen,[15] and 0.7% from Hurley and Maclaren.[16]

The differences seen between the curves from the various studies could have been due to a multitude of factors apart from random variation. These factors include the nature of the study group (longitudinal/cross-sectional), the period of study, the presence of miner selection effects, the type of x-ray reading performed, the x-ray readers, the x-ray classification used, and the methods of model fitting and extrapolation. Considering the potential for variation, the agreement between the various studies is noteworthy.

Each study has its stronger and weaker features in relation to the others. For instance, the latest analysis, by Hurley and Maclaren,[16] is based on the study of a very large group of coal miners for whom health-related selection effects were probably small. However, the very size of the group precluded the rereading of the films using a tightly controlled protocol. For the ISS group, the opposite was true: the miners were probably fairly highly selected, having to be predominantly face workers over a 10-year period to be included in the study, but the quality of the x-ray readings was high because the x-rays were all specially reread by eight readers.

Results from successive British analyses have suggested an increasing risk of CWP over that which seemed apparent when the 1969 federal law was enacted. Among the many possible factors that could be responsible for this increase, two are discussed further. The first involves the choice of models. This applies to the ISS analysis compared with the later studies. That initial investigation used a form of model that permitted the estimation of a threshold. In contrast, all later analyses have assumed a continuum of response down to zero dust exposure.

TABLE 1—Predicted Percentages of Miners with CWP by Age

CWP Category	Age at Start of 5-Year Interval							
	23	28	33	38	43	48	53	58
0	99.96	99.71	99.15	98.27	97.05	95.46	93.49	91.15
1	0.04	0.27	0.75	1.46	2.40	3.55	4.89	6.41
2	0.00	0.02	0.09	0.22	0.45	0.76	1.15	1.60
3	0.00	0.00	0.00	0.01	0.02	0.05	0.09	0.14
PMF	0.00	0.00	0.01	0.03	0.08	0.18	0.37	0.70

Note. These predictions assume miner starts work at 18, works continually in a 2 mg/m^3 environment, obtains an x-ray every 5 years, and transfers to a low-dust (1 mg/m^3) job immediately after category 1/0 or greater is seen. Predictions are derived from results of Hurley and MacLaren.[16]

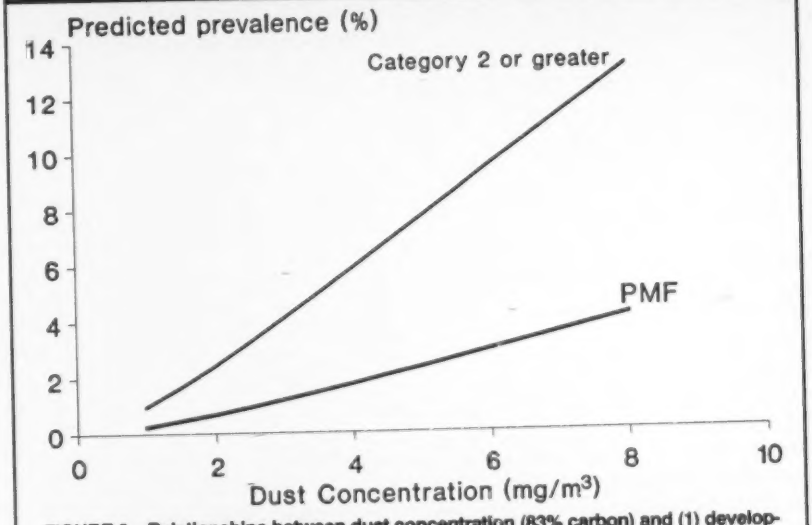


FIGURE 3—Relationships between dust concentration (83% carbon) and (1) development of category 2 or greater simple coal workers' pneumoconiosis, and (2) progressive massive fibrosis over 40 years (derived by using results given in Hurley and Maclaren[16]).

The second factor, which relates only to PMF, lies in the incorporation of information in the later studies that was ignored in the setting of the federal level: that is, incidence of PMF from simple CWP categories 0 and 1. In an early influential study of PMF incidence,[7] the risk of PMF from categories 0 and 1 appeared low compared with the obviously greater risk from the higher categories of simple CWP. For this reason, emphasis on prevention of category 2 became the major focus of the strategy for prevention of PMF. However, results from later studies of PMF incidence show a much lower gradient in risk across the categories of simple CWP, with increased risks for categories 0 and 1, and smaller risks for categories 2 and 3.[11–13] Furthermore, as exemplified in the findings of Shennan et al.,[12] the fact that most miners have category 0 or 1 simple CWP implies that a substantial fraction of the total number of PMF cases would arise from those categories, even though the risk for an individual miner might be small in that area. The later exposure-response studies with PMF as a response variable took into account this route of contracting PMF and, as a consequence, imply a higher overall risk.

Random variation obviously played a role in causing differences between the results from the various studies. However, only one of the studies provides estimates of uncertainty in the predictions. This was the ISS analysis, for which approximate confidence bounds for the 35-year predictions of risk were given. These actually indicate a range quite similar to that seen between the various PFR studies noted above.

Although the methods and models used in the PFR analysis have varied, the initial and latest studies used similar approaches to estimate the risks of CWP over a working life. In the recent Hurley and Maclaren analysis,[16] 5-year incidence and progression data were compounded into 40-year predictions. This technique is very similar to that used in the ISS, in which 10-year probabilities were compounded into 35-year estimated risks. Evidence that the ISS approach was reasonable was seen in the analysis of Hurley et al.,[17] which derived 35-year estimates of incidence directly without compounding and obtained very similar predictions to those of the ISS study (Figure 2). Hurley and Maclaren[16] expressed preference for the compounding method, since it employs the most important factor (initial category) for prediction of category 5 years later. They considered that the predictions derived using the compounding method[16] were more reliable than earlier results obtained using other methods.[15]

Application of British Results to US Mines and Miners

The validity of extrapolating results from previous studies of British mines and miners to the US situation is not known. Moreover, it may never be possible to determine exactly how valid the British findings are for the United States, given that such an evaluation would require knowledge that is now unavailable (such as that on particle-size distributions or composition for mines that are now closed). However, the following text discusses some of the more important aspects that should be considered.

Environmental factors of relevance encompass the nature of the coal mined and its surrounding or intermixed rock, coal mining practices, and other mine-specific variables. Hurley and colleagues have cautioned against indiscriminate use of their results in this respect: "There are several reasons, however, why the estimates might be inexact or misleading when applied to other conditions. . . . It is clear that at some collieries risks are far lower, or far higher, than the average values. This implies that our probability estimates will not necessarily reflect the risks to coal miners generally (in Britain or elsewhere) unless the dust concentrations and sources of unexplained colliery variability occur in a pattern broadly similar to that observed in this study."[17]

There are, however, several points of similarity between British and US mining conditions. The rank of coal mined in the two countries appears to vary over a similar range, from anthracite to high volatile bituminous.[18,19] Moreover, the average level of quartz in US mines currently also appears to be similar to that seen in British mines. Data from a recent analysis of quartz in US mines suggest that the overall mean quartz level is around 5%.[20] This compares with a mean mine average of about 4% reported for the PFR mines,[18] although this comparison does not take into account differences in the methodology used to estimate quartz levels.

Although the coals may be similar in general, there are some important differences in mining methods. Coal has been mined for many years in Britain using the longwall extraction system, a method that is relatively new to the United States. Mine operators in this country have relied on continuous miners and other coal-getting methods. Because the various methods may give rise to different particle-size distributions and involve cutting into roof, floor, or dirt band rock to differing degrees, the resulting dust clouds may vary considerably in type and composition and thus may have differing fibrogenic potential.

Another factor that needs to be considered relates to dust level estimation methods and procedures. Currently, these are different in the United States compared with those used in the British exposure-response studies. The latter involved the collection of dust concentration information, first using the thermal precipitator and then using gravimetric methods based on an instrument with a horizontal elutriator (the British Mining Research Establishment [MRE] dust sampler).[21] Later, the thermal precipitator data were converted into gravimetric equivalents using the results of comparative side-by-side sampling exercises. Furthermore, the British data were gathered according to a strategy designed specifically for epidemiological research. US procedures, in contrast, are somewhat different, being designed to assess compliance with the dust standards. Moreover, the samples are obtained using a cyclone instrument, the results of which are multiplied by a factor of 1.38 to make them equivalent to dust concentrations obtained using the MRE sampler. The uncertainties inherent in these different methods and procedures need to be considered when interpolating British exposure-response curves for application to the US coal mining situation.

Certain factors related to x-ray reading also need to be considered. First, there is evidence that US readers have tended to report more abnormality than their British colleagues in the past.[22] If this tendency still exists, prevalences of CWP estimated for US coal miners based on x-ray interpretations by US readers will be greater than those expected from the British exposure-response curves. Second, the results from the British studies summarized here were based on x-ray readings from international classification standards for pneumoconiosis that are no longer current. Because the classification is revised periodically with the introduction of different reference films, this change could easily affect the amount of abnormality reported in the future. Third, it must be noted that virtually all the PFR analyses have concentrated on small rounded opacities, whereas general US practice is to refer to combined opacity scores (rounded plus irregular opacity profusions). Moreover, the latest international classification[23] does not permit the independent assessment of rounded opacities in the same manner as past classifications. These factors imply that care will have to be taken in comparing future observed prevalences for US miners with predicted levels derived from studies using British readers and based on differing x-ray classification schemes and reference films.

Because conditions other than CWP can lead to the presence of x-ray markings that mimic simple CWP, it might be expected that a non-zero prevalence of abnormality would be seen in any cohort of subjects, including nonexposed workers. Moreover, this baseline prevalence would increase with age. Even though age is included in some of the British analyses, the form of the model forces the predictions to zero at zero dust. Hence, to date, this phenomenon has not been fully explored or incorporated into any existing analyses.

By law, current dust exposures for US underground coal miners must be less than 2 mg/m^3. This range of dust levels is actually at the extreme lower end of the range studied in the ISS analysis and other PFR analyses. Moreover, no systematic exploration of the form of the exposure-response curve in this region of dust exposure has been undertaken; hence, the proper form of this curve has not been established exactly. In particular, the existence of a threshold level of dust exposure below which no disease will occur cannot be ruled out.

As noted earlier, only the ISS analysis permitted a threshold to be estimated. There, an arc-sine transformation was fitted to the proportion of miners with cat-

egory 2 or greater simple CWP. This form of exposure-response curve is sigmoid shaped and resembles the probit and logistic transforms. It differs from them, however, in that the curve can touch the exposure axis at any point. In fact, the curve did touch the axis at dust concentration close to 2 mg/m^3, which may indicate that a threshold exists at that point. However, because no other form of model was fitted, it is not possible to determine whether the arc-sine curve was the most appropriate.

Although caution should be used in applying the British findings to the US situation, some evidence exists supporting their use. In 1984, results of a 9-year study of incidence of CWP in US coal miners were published.[24] These showed that the amount of incidence and progression of the disease was consistent with predicted levels from the initial British investigation,[9] given dust exposures mandated by federal law over the study period. These findings were supported by a study of new miners over a similar period.[25] However, these two studies covered too short a period for reliable long-term indications; further study is planned and in progress.

Evidence of a continuing trend toward lower prevalences of CWP in US miners has been reported recently.[2,3] Will the current federal disease prevention measures ensure that this trend continues until CWP is eventually eliminated from the industry? Evidence from the British studies suggests not, although the uncertainties inherent in extrapolation make it difficult to be sure as to the exact level and severity of disease to be expected. It should also be borne in mind that many miners do not transfer to a low-dust job after having an x-ray positive for CWP,[26] and that overexposure may be occurring.[27] Both phenomena suggest that observed levels of CWP incidence will be higher than those predicted. Finally, it should be noted that the results presented here are restricted to CWP; the potential for nonpneumoconiotic occupational lung diseases in coal miners should not be overlooked. □

Acknowledgments

Thanks are due to Barbara Bonnett for help in preparation of the figures and in other programming, and to the many who have helped in preparing and reviewing the manuscript.

References

1. Coal Mine Health and Safety Act. US Pub L No. 91–173, Stat 2917.
2. Attfield MD, Althouse RB. Surveillance data on US coal miners' pneumoconiosis, 1970 to 1986. *Am J Public Health.* 1992; 82:971–977.
3. Attfield MD, Castellan RM. Epidemiological data on US coal miners' pneumoconiosis, 1960 to 1988. *Am J Public Health.* 1992;82:964–970.
4. Kerr L. Black lung. *J Public Health Policy.* 1980;1:50–63.
5. Parkes RW. *Occupational Lung Disorders.* Boston, Mass: Butterworths; 1982: 175–232.
6. Collins HPR, Dick JA, Bennett JG, et al. Irregularly shaped small shadows on chest radiographs, dust exposure, and lung function in coalworkers' pneumoconiosis. *Br J Ind Med.* 1988;45:43–55.
7. Cochrane AL. The attack rate of progressive massive fibrosis. *Br J Ind Med.* 1962; 19:52–64.
8. Fay JWJ, Rae S. The pneumoconiosis field research of the National Coal Board. *Ann Occup Hyg.* 1959;1:149–161.
9. Jacobsen M, Rae S, Walton WH, Rogan JM. The relation between pneumoconiosis and dust exposure in British coal mines. In: Walton WH, ed. *Inhaled Particles III.* Old Woking, England: Unwin Brothers; 1971: 903–919.
10. Jacobsen M. Progression of coal workers' pneumoconiosis in Britain in relation to environmental conditions underground. In: *Proceedings of the Conference on Technical Measures of Dust Prevention and Suppression in Mines, Luxembourg, 11–13 October 1972.* Luxembourg: Commission of the European Communities; 1973:77–93.
11. McLintock JS, Rae S, Jacobsen M. The attack rate of progressive massive fibrosis in British miners. In: Walton WH, ed. *Inhaled Particles III.* Old Woking, England: Unwin Brothers; 1971:933–952.
12. Shennan DH, Washington JS, Thomas DJ, et al. Factors predisposing to the development of progressive massive fibrosis. *Br J Ind Med.* 1981;38:321–326.
13. Hurley JF, Maclaren WM, Alexander WP, et al. Factors influencing the occurrence of progressive massive fibrosis in British coalminers. Report no. TM/84/2. Edinburgh, Scotland: Institute of Occupational Medicine; 1984.
14. Hurley JF, Alexander WP, Hazledine DJ, et al. Exposure to respirable coalmine dust and incidence of progressive massive fibrosis. *Br J Ind Med.* 1987;44:661–672.
15. Hurley JF, Jacobsen M. Occupational hygiene implications of new results on progressive massive fibrosis in working coalminers. *Am Conf Gov Ind Hyg.* 1986;14: 85–90.
16. Hurley JF, Maclaren WM. Dust related risks of radiological changes in coalminers over a 40-year life: report on work commissioned by NIOSH. Report no. TM/87/ 09. Edinburgh, Scotland: Institute of Occupational Medicine; 1987.
17. Hurley JF, Burns J, Copland L, et al. Coalworkers' simple pneumoconiosis and exposure to dust at 10 British coalmines. *Br J Ind Med* 1982;39:120–127.
18. Walton WH, Dodgson J, Hadden GG, et al. The effect of quartz and other non-coal dusts in coalworkers' pneumoconiosis. In: Walton WH, ed. *Inhaled Particles IV.* Oxford, England: Pergamon Press; 1977:669–689.
19. Given PH. An essay on the organic geochemistry of coal. In: Gobarty ML, Larsen JW, Wender I, ed. *Coal Science.* Vol 3. New York, NY: Academic Press; 1984:70–86.
20. Tomb TF, Peluso RG, Parobeck PS. Quartz in United States coal mines. *Am Conf Gov Ind Hyg.* 1986;14:513–519.
21. Dunmore JH, Hamilton RJ, Smith DSG. An instrument for the sampling of respirable dust for subsequent gravimetric assessment. *J Sci Instrum.* 1964;41:669–672.
22. Reger RB, Amandus HE, Morgan WKC. On the diagnosis of coalworkers' pneumoconiosis: Anglo-American disharmony. *Am Rev Respir Dis.* 1973;108:1186–1191.
23. International Labour Office. *International Classification of Radiographs of Pneumoconiosis.* Occupational Safety and Health, Series No. 22 (Rev 80). Geneva, Switzerland: International Labor Office; 1980.
24. Attfield MD, Reger RB, Glenn RE. The incidence and progression of pneumoconiosis over nine years in US coal miners: I. principal findings. *Am J Ind Med.* 1984;6: 407–415.
25. Althouse RB, Attfield MD, Kellie S. Use of data from the x-ray surveillance program for underground coal workers to evaluate the efficacy of the dust standard, presented at Conference on Medical Screening and Biological Monitoring for the Effects of Exposure in the Workplace, Cincinnati, July 1984. *J Occup Med.* 1986;28:741–745.
26. Hoffman JM. X-ray surveillance and miner transfer programs: efforts to prevent progression of coal workers' pneumoconiosis. *Ann Am Gov Ind Hyg.* 1986;14:293–297.
27. Mine Health and Safety Administration. Peabody Coal Company pleads guilty to dust sample tampering. Mine Safety and Health Administration; January 17, 1991. Washington DC: Press Release USDL 91–23.

Job Strain and Prevalence of Hypertension in a Biracial Population of Urban Bus Drivers

Cheryl L. Albright, PhD, MPH, Marilyn A. Winkleby, PhD, MPH, David R. Ragland, PhD, June Fisher, MD, and S. Leonard Syme, PhD

ABSTRACT

Objectives. In this study we tested the association between occupational stress—as measured by job demands, decision latitude, and job strain—and hypertension in a population of 1396 Black and White bus drivers.

Methods. Height, weight, blood pressure, and medical history were assessed by physical exam. Drivers completed a questionnaire assessing their work schedules, personal habits, and self-perceptions about job demands and decision-latitude.

Results. Univariate analyses revealed significant inverse associations; lower levels of job demands and job strain were associated with a higher prevalence of hypertension for Blacks and Whites. After 12 confounding variables were controlled for, the association between these two measures of occupational stress and hypertension became nonsignificant. Decision latitude was also not significantly associated with hypertension.

Conclusions. Our findings are inconsistent with previous studies' findings of a positive association between job strain and chronic diseases. The difference in results may be explained by our incorporation of individuals' perceptions in the measurement of occupational stressors and our use of individuals from a single occupation with comparable job responsibilities and income, thus controlling for potential confounding by social class. (*Am J Public Health.* 1992;82:984–989)

Introduction

Since the 1930s, social scientists have conducted research to examine how psychosocial stressors such as life events,[1,2] daily hassles,[3] behavioral characteristics,[4–6] and stressful working conditions[7–12] may be associated with the incidence and prevalence of various chronic diseases. Particular interest has been focused on the relationship between hypertension and/or coronary heart disease and occupational stressors, operationalized as specific job problems, high job demands, job dissatisfaction, and low job rewards.[13–15] This work has yielded inconsistent findings, with some studies showing positive associations[13,14,16–25] and others indicating null or negative associations.[6,12,15,26–29]

Work by Karasek and his colleagues suggests that there is a strong association between job strain, conceptualized as the joint effect of high job demands and low decision latitude or control, and an increased prevalence and/or incidence of coronary heart disease and myocardial infarction when traditional risk factors such as age, race, cholesterol, and smoking are controlled for.[14,18,20] Also, job strain has recently been shown to be positively associated with elevated diastolic blood pressure and a higher left-ventricular-mass index in a case-control study of employees from several occupations (odds ratio of 3.1).[25]

Although some studies have used case-control or prospective designs,[18,25] most of the studies using Karasek's job-strain variable have employed an ecologic study design.[14,15,20,21,30–32] In these studies an imputation method assigns job-strain scores to individuals within the same occupational census code (data obtained from large surveys such as the Health Examination Survey and the Health and Nutrition Examination Survey), then compares rates of myocardial infarction, for example, between high- and low-strain occupations.[14,20,21,30] This methodology does not permit evaluations of the associations between workers' self-perceptions of their jobs and disease status. In addition, it does not allow assessment of whether workers performing the same job duties have different perceptions of their job experience and different disease prevalence rates.

The present study was designed to test the association between self-reported job strain and clinically diagnosed hypertension within a population of bus drivers. Bus driving is an occupation characterized as having a high level of job strain as well as high rates of hypertension.[33,34]

We hypothesized that drivers who perceived their job as demanding, with little opportunity for control or decision latitude, would exhibit a higher prevalence of hypertension than would drivers who did not perceive their job in this way. In addition, we hypothesized that the association would persist after we controlled for other biobehavioral risk factors related to hypertension, such as age, race, sex, ed-

Cheryl L. Albright and Marilyn A. Winkleby are with the Stanford Center for Research in Disease Prevention, Palo Alto, Calif. David R. Ragland and S. Leonard Syme are with the School of Public Health, University of California, Berkeley, Calif. June Fisher is with the Center for Municipal Occupational Safety and Health, San Francisco General Hospital, and the University of California at San Francisco.

Requests for reprints should be sent to Cheryl L. Albright, PhD, MPH, Stanford University School of Medicine, 1000 Welch Road, Palo Alto, CA 94304-0146.

This paper was submitted to the Journal July 10, 1991, and accepted with revisions December 10, 1991.

ucation, body mass index, family history of heart disease, and alcohol intake.

Methods

The present analysis was conducted as part of a larger occupational health study of urban bus drivers employed by the San Francisco Municipal Railway system. Data collection was incorporated into an existing system of medical examinations conducted as part of the drivers' required biannual license renewal process. Drivers who were examined from December 1983 to September 1985 were eligible for recruitment into the study.

Study Population

The target population consisted of 1845 San Francisco bus drivers, all of whom completed a medical examination (28 with nonstandardized blood pressure data were excluded). Seventy-seven percent completed the optional occupational psychosocial questionnaire, yielding a final study population of 1396 drivers for whom both the medical examination and optional questionnaire data were complete. A previous comparison of the drivers who completed the questionnaire with those who did not revealed no statistically significant differences between respondents and nonrespondents in prevalence of hypertension and selected demographic and behavioral characteristics.[26]

Measurement of Variables

Detailed descriptions of the data collection methods and medical procedures have been published previously.[26] Briefly, the medical examination included assessments of height, weight, blood pressure, medical history, and personal health habits. All physiological assessments were obtained in a clinic setting during normal working hours (i.e., 8 AM to 5 PM). Three standardized blood pressure measurements were taken by physicians and nurses according to the protocol defined by the Hypertension Detection and Follow-up Program.[35] All readings were obtained from the right arm with standard mercury manometers while the participants were seated. Drivers were classified as hypertensive if (1) the mean of their second and third systolic readings was equal to or greater than 140 mm Hg, (2) the mean of the second and third diastolic readings was equal or greater than 90 mm Hg, or (3) they were currently using antihypertensive medication.

After the medical examination, drivers completed a questionnaire assessing their work schedules, personal habits (health and job-related), self-perceptions about specific job problems,[26] and the 37-item job-content instrument developed by Karasek.[36,37] The present analysis focused on the three composite scales assessing psychological job demands, decision latitude, and job strain, the latter being a composite scale that represents the interaction between demands and decision latitude.[14,18,30] For each question in the scales, respondents rated (on a scale of one to four) the degree to which they agreed or disagreed with statements describing their job or job environment. A high score represents a high level of perceived job demands and/or a high level of perceived decision latitude over the job.

Measurement of Covariables

Biobehavioral factors potentially related to hypertension and assessed as confounding factors included the following: (1) age, (2) martial status, (3) education, (4) income, (5) number of years employed as a bus driver, (6) race, (7) body mass index, (8) family history of heart disease, (9) smoking (pack years), (10) fitness, (11) alcohol consumption in a week, and (12) caffeine intake per day (see Appendix A for definitions of covariables).

Statistical Analyses

Logistic analyses were conducted with hypertension status as the outcome variable. Univariate analyses were performed initially, followed by multivariate logistic analyses that included job demands and decision latitude as the main effects, the job strain variable as the interaction term, and the potential 12 confounding factors. Job demands and decision latitude were tested separately as continuous variables and as dichotomous "median split" variables with high job demand > 31.5 and low decision latitude ≤ 30.9. These two categorical variables were used to create the four-level rank-ordered job-strain variable that corresponded to Karasek's 2×2 model of job strain, with low job demand and low decision latitude representing the passive job category; low job demand and high decision latitude representing the low-strain category; high job demand and high decision latitude representing the active category; and high job demand and low decision latitude representing the high-strain category. The latter category was hypothesized to carry the greatest risk.[14,20,25]

Results

The sample was predominantly male (91%) and was 60% Black, 17% White, 13% Asian, and 8% Hispanic. The drivers' ages ranged from 20 to 65 years; the mean age was 42. The majority (63%) had been employed by the San Francisco bus company for 9 years or longer.

The mean and standard deviation was 32.0 ± 6.04 (range, 12 to 48) for the job demands scale and 30.8 ± 5.00 (range, 14 to 45) for the decision latitude scale. The mean value for job demands is similar to those reported previously, but the mean value for decision latitude is somewhat lower[15,25,37] (see Appendix B).

Table 1 presents the descriptive statistics for the drivers according to job strain category. The active job strain category contained the lowest proportion of hypertensives. Chi-square analyses were significant for these descriptive variables; higher percentages of younger drivers, women, and drivers with less than 10 years of employment fell into the active and high-strain categories. (When we examined social-class variables, our correlational analyses revealed a positive association between income and hypertension [$r = .07$, $P < .02$], but a negative association between education and hypertension [$r = -.060$, $P < .02$].)

Table 2 shows the percentage of hypertensives within each quintile of job demands and decision latitude. According to our hypothesis, hypertension rates should be higher within the higher levels of job demands and lower within the higher levels of decision latitude. Although not statistically significant, the trend for decision latitude was in the hypothesized direction. Job demands yielded a significant chi-square with a P value of less than .01; however, the trend was the opposite of the hypothesized relationship. Analyses by race revealed similar trends in prevalence of hypertension across the job-demands quintiles for Black and White drivers (sample sizes for Asian and Hispanics drivers were insufficient for analyses).

Because drivers with more seniority are allowed to select their bus routes (i.e., they have more latitude to select less demanding routes), it is possible that older drivers with more hypertension might perceive fewer job demands. This possibility is supported by the data reported in Table 1. Of those drivers in the passive category, 24% were older than 50, whereas of those in the active or high-strain category, only 12% were older than

TABLE 1—Characteristics of Respondents by Job-Strain Category (%)

	Job-Strain Category					
	Passive (n = 295)	Low Strain (n = 409)	Active (n = 357)	High Strain (n = 345)	χ^2	Significance
Age						
20–29 y	3	2	3	2	42.2 (df = 12)	P < .0001
30–39 y	34	38	50	48		
40–49 y	42	40	35	38		
50–59 y	20	17	11	10		
60–69 y	4	3	1	2		
Male	94	94	88	88	14.4 (df = 3)	P < .002
Race						
Black	62	57	61	62	30.9 (df = 12)	P < .002
White	18	14	15	20		
Asian	11	14	14	12		
Hispanic	6	13	9	5		
Other	3	2	1	1		
Years employed						
< 4	28	34	38	25	78.9 (df = 21)	P < .0001
5–9	22	26	30	37		
10–14	18	16	16	23		
15–19	14	12	10	9		
20–24	13	7	5	3		
25–29	3	4	1	2		
30–34	1	2	0	3		
> 35	1	2	1	0		
Hypertensive, %	36	35	25	31	12.7 (df = 3)	P < .005

TABLE 2—Prevalence of Hypertension (% Hypertensive) by Quintiles of Job Demands and Decision Latitude

Quintile	Job Demands[a]	Decision Latitude[b]
1 (Lowest)	39	36
2	36	32
3	29	31
4	30	31
5 (Highest)	27	29

[a] $\chi^2 = 12.7$ (df = 4), $P < .01$.
[b] $\chi^2 = 2.7$ (df = 4), $P = .61$.

50. Thus, a multivariate logistic analysis was needed to investigate this association both by controlling for age and length of employment and to test any interaction between job strain and length of employment.

As shown in Table 3, univariate analyses revealed significant inverse associations between the prevalence of hypertension and job demands and job strain (odds ratios of .84 and .88, respectively). The univariate test for decision latitude was in the hypothesized inverse direction but was not statistically significant. Multivariate logistic regression analyses in which age, years of employment, body mass index, and nine other factors were controlled for revealed no significant associations between any of the three occupational stressors and the prevalence of hypertension. Similar results were found for both the continuous and dichotomous quantifications of the three variables.

To examine whether an interaction existed between perceived job strain and covariables such as age, sex, and race, we assessed evidence of a first-order multiplicative interaction, using unconditional maximum-likelihood estimation methods.[38] Product interaction terms (between each of the covariables and job strain) were added to the multivariate model and the model was compared with a reduced model without interaction terms. A chi-square statistic, used to test the difference between the two models' log likelihood ratios, revealed no significant interaction between the job stress variables and any of the covariables.

Separate multiple regression analyses examining the impact of the three occupational stress indices on mean systolic and mean diastolic blood pressures within the subsample of normotensives (systolic blood pressure < 140 and diastolic blood pressure < 90) were also not statistically significant.

Discussion

We found univariate inverse relationships between job demands, job strain, and the prevalence of hypertension in a biracial blue-collar population of urban bus drivers. These negative associations were opposite to our a priori hypothesis of a positive relationship between these variables. The univariate inverse associations were not, however, supported by the multivariate logistic analysis that adjusted for 12 covariables. Thus, in the present study we found no association between subjective reports of occupational stress and prevalence of hypertension.

Previous studies using subjective self-reported measures of occupational stress have yielded inconsistent results. Some have yielded positive associations with diastolic or systolic blood pressure,[13,16,22,25] whereas others have found no associations with elevated blood pressure or hypertension.[6,12,29,32] These inconsistencies in the literature could be related to differences in populations, study design, job stress instruments, and/or outcome measures. However, the similarities between our study and the recent case-control investigation of job strain, diastolic blood pressure, and left-ventricular-mass index conducted by Schnall et al.[25] warrants discussion. Both studies used the same instruments to assess job strain and similar statistical analyses. Schnall et al. found that individuals from different occupations with high-strain jobs had an increased likelihood of having elevated diastolic blood pressure, whereas our analyses revealed that, within one occupation, there was no significant association between job strain and prevalence of hypertension or systolic/diastolic blood pressure. Although there were methodological differences between these two studies (e.g., selection of subjects and blood pressure measures), two major differences between our study and that of Schnall et al. require elaboration. Our study population was largely Black (60%), whereas that of Schnall et al. was largely White (80%). Schnall and colleagues examined job strain across a variety of occupations, whereas our study included only bus drivers. Previous analyses of bus drivers' work have shown that drivers are exposed to a wide range of psychological as well as physical stressors,[33] and comparisons across occupations have shown bus drivers to have higher rates of hypertension and other diseases than do workers in other blue-collar occupations.[33,34]

One explanation for the divergent findings of these two studies is that our study population, which was from one occupational group, could have exhibited insufficient variability in the assessments of job strain. We found, however, that the mean value for job demands for our study population was quite similar to that reported by Reed et al. and Karasek et al. in cross-occupational studies[15,37] and by Schnall et al. in their recent case-control study[25] (the mean value for decision latitude was somewhat lower in the present study; see Appendix B). Although our standard deviation for decision latitude was considerably lower than those reported by Karasek et al.,[37] it was similar to those reported by Reed et al.[15] and Schnall et al.[25]; thus, the within-occupation variability of self-reported job stress in the present study is similar to the between-occupation variability found in the latter two studies. This similarity indicates that individuals working in jobs with similar duties can have different perceptions of their duties and can cope with them with varying degrees of success. And although the duties may be similar, there may be many factors that vary between workers (e.g., for bus drivers, the number of passengers or the traffic volume on different bus routes). Thus there does not appear to be bias from limited variability in our measure of occupational stress.

Another reason for the inconsistency between the present study and Karasek's earlier studies may be the different socioeconomic status (SES) of the respondents. In the analyses of Karasek and his coworkers, education is often used as a covariable[14,18,20,25,32]; however, factors other than education that are related to SES (e.g., income, job status) may confound the association between occupational stress and chronic disease. For example, Marmot and his coworkers found that low SES, as measured by grade of employment, was associated with increased mean blood pressure, prevalence of hypertension, and stroke mortality.[39,40] Several similar investigations in which occupation has been used as an indicator of SES have found that blood pressure rises as SES decreases.[41,42] In contrast to the work of Karasek and his colleagues, the present study involved a population that was similar in socioeconomic status, and individual differences in education and income were controlled for in the analysis. This method of controlling for SES in both the design and analytic phases of the study may explain some of the disparities in results.

TABLE 3—Logistic Analyses of Hypertension and Occupational Stressors

	Regression Coefficient	Odds Ratio (95% CI)	*P* Value
Univariate analyses			
Job demands	−.028	0.84 (.75, .94)	<.01
Decision latitude	−.018	0.91 (.82, 1.02)	.12
Job strain[a]	−.119	0.88 (.78, .98)	.03
Multivariate analyses[b]			
Job demands	−.008	0.95 (.84, 1.03)	.46
Decision latitude	−.017	0.92 (.81, 1.04)	.19
Job strain[a]	−.011	0.98 (.86, 1.13)	.85

[a]Job strain was measured by a four-level rank-ordered combination of job demands and decision latitude.
[b]Multivariate analyses tested the association between occupational stressors and prevalence of hypertension with adjustment for age, race, sex, education, income, marital status, years employed, body mass index, smoking, caffeine, alcohol, and fitness.

Differences between the present study and previous studies are unlikely to be due to inadequate sample size. Of the 1407 subjects in our sample, 448 were classified as hypertensive. This number is more than the total sample size used in two similar studies of occupational stress and diastolic blood pressure (n = 288 and n = 215).[13,25] Thus, compared with two studies that found a positive association between occupational stress and diastolic blood pressure, our study had almost six times as many subjects and five times as many hypertensives.

Inconsistencies could have, however, occurred because of selection or measurement bias. Selection bias could have occurred if the subjects in the present study were atypical with regard to their perceptions about their job or their risk factors for hypertension, or if those hypertensives with the highest level of job stress were more likely to terminate employment. As discussed above, the mean values for the occupational stress variables were quite similar to those reported in previous studies; thus the drivers' job scores were not unusual. Furthermore, our study population and that used by Schnall et al. were extremely similar in terms of risk factors for hypertension, such as age, body mass index, prevalence of smokers, and systolic and diastolic blood pressures. Finally, in a previous study of the same sample in which drivers were asked to rate specific job stressors, we found no significant differences between the job-stress scores of hypertensives who left the job and the scores of those who remained.[26] Although measurement bias is possible, it was minimized by (1) distinct separation, both in time and in setting, of blood pressure measures and driver's ratings of job strain; (2) fidelity to the Karasek instrument; and (3) strict adherence to standardized protocols for measuring blood pressure.

Therefore, these potential biases do not appear to account for our negative findings. Bus drivers' self-perceptions about the level of job demand and decision latitude involved in their job does not appear to be associated with hypertension or high blood pressure. In fact, there may even be an inverse association between job stress and hypertension. Significant inverse relationships between subjectively assessed job stressors and hypertension or blood pressure were recently found in two studies.[26,32] In the first study,[26] Winkleby et al. found an inverse association between specific work-related stressors (e.g., accidents, traffic, and tight schedules) and hypertension (odds ratio = 0.84; CI = 0.75, 0.99; $P < .04$), using the same sample of bus drivers as the current study. In the second study an inverse association between job demands and systolic blood pressure (regression coefficient = −.79; $P < .05$) was found in one of the five databases reviewed in a meta-analysis by Pieper et al.[32] These studies suggest that low levels of job strain may be associated with higher hypertension or that hypertensives may suppress their feelings toward job stressors by underreporting the frequency of specific stressful situations and the levels of emotional reactivity associated with these events.[26]

Summary and Implications

In summary, we studied workers with comparable job responsibilities and similar income levels in an objectively defined high-stress occupation and found no significant association between hypertension and job demands, decision latitude,

or job strain. This lack of association is in contrast to the findings of Karasek et al., who examined workers from a wide range of social classes. If the studies of Karasek et al. did not adequately control for SES, then it may *not* be high demand/low decision latitude that is associated with high coronary heart disease morbidity or mortality, but the low SES that accompanies occupations with such stressors. We hypothesize that social class could confound the association between occupational stress and chronic disease, or that SES and occupation may be so highly correlated that it is difficult to separate their independent effects. Therefore, until the effects of SES can be distinguished more precisely from occupation-based indicators of job stress, ecologic investigations of different occupations should be interpreted cautiously.

Our results raise questions about the relationship between Karasek and his colleagues' operationalization of occupational stress and the prevalence of hypertension. In addition, the interrelationships between coping strategies (denial and repression), occupational stress, social class, and chronic disease remain unclear. Future investigations should rigorously evaluate coping styles, how they interact with occupational stressors, and whether they precede or follow the development of hypertension.

To fully address the inconsistencies evidenced in this study, future studies should use prospective designs that assess objective as well as subjective measures of occupational stress; evaluate coping styles (especially repressive coping styles); control for the various dimensions of social class; and follow workers for the development of multiple chronic diseases such as hypertension, coronary heart disease, and stroke. Obviously this type of study would be expensive and difficult to implement, but smaller studies investigating how hypertensives explicitly deal with various occupational stressors, measured both subjectively and objectively, may provide insights into the inverse and null associations found between occupational stressors and hypertension. □

Acknowledgments

This work was supported in part by the Urban Mass Transit Authority of the United States Government Department of Transportation under grant CA-06-0160. Dr Albright was supported by US Public Health Service Training Grant T32 HL07365 from the National Heart, Lung, and Blood Institute, Dr S. Leonard Syme, director.

References

1. Dohrenwend BS, Dohrenwend BP, eds. *Stressful Life Events: Their Nature and Effects.* New York, NY: John Wiley & Sons; 1974.
2. Rahe RM, Arthur RH. Life change and illness studies. *J Human Stress.* 1978;4:3–15.
3. DeLongis A, Coyne JC, Dakof G, Folkman S, Lazarus RS. Relationship of daily hassles, uplifts, and major life events to health status. *Health Psychol.* 1982;1:119–136.
4. Syme SL. Social determinants of health and disease. In: Last JM, ed. *Maxcy-Rosenau: Public Health and Preventive Medicine.* 12th ed. Norwalk, Conn: Appleton Century Crofts; 1986:953–970.
5. Matthews KA, Weiss S, Detre T, et al, eds. *Handbook of Stress, Reactivity and Cardiovascular Disease.* New York, NY: John Wiley & Sons; 1986.
6. Haynes SG, Levine S, Scotch N, Feinleib M, Kannel WB. The relationship of psychosocial factors to coronary heart disease in the Framingham Study. I. Methods and risk factors. *Am J Epidemiol.* 1978;107:362–383.
7. Buring JE, Evans DA, Fiore M, Rosner B, Hennekens CH. Occupation and risk of death from coronary heart disease. *JAMA.* 1987;258:791–792.
8. World Health Organization Expert Committee. *Identification and control of work-related diseases.* Geneva, Switzerland: World Health Organization; 1985. World Health Organization Technical Report Series 714.
9. Cooper CL, Payne R, eds. *Stress at Work.* New York, NY: John Wiley & Sons; 1978.
10. Holt RR. Occupational stress. In: Goldberger L, Breznitz S, eds. *Handbook of Stress: Theoretical and Clinical Aspects.* New York, NY: Free Press; 1982:419–444.
11. *Psychosocial Factors at Work: Recognition and Control.* Geneva, Switzerland: International Labour Office; 1986. Report of the Joint ILO/WHO Committee on Occupational Health.
12. Caplan RD, Cobb S, French JRP, et al. *Job Demands and Worker Health: Main Effects and Occupational Differences.* Ann Arbor, Mich: Institute for Social Research; 1980.
13. Matthews KA, Cottington EM, Talbott E, Kuller LH, Siegel JM. Stressful work conditions and diastolic blood pressure among blue collar factory workers. *Am J Epidemiol.* 1987;126:280–291.
14. Karasek RA, Theorell T, Schwartz JE, Schnall PL, Pieper CF, Michela JL. Job characteristics in relation to the prevalence of myocardial infarction in the US health examination survey (HES) and the health and nutrition examination survey (HANES). *Am J Public Health.* 1988;78:910–918.
15. Reed DM, LaCroix AZ, Karasek RA, Miller D, MacLean CA. Occupational strain and the incidence of coronary heart disease. *Am J Epidemiol.* 1989;129:495–502.
16. Van Dijkhuizen N, Reiche H. Psychosocial stress in industry: a heartache for middle management? *Psychother Psychosom.* 1980;34:124–134.
17. Aro S. Occupational stress, health-related behavior, and blood pressure: a 5-year follow-up. *Prev Med.* 1984;12:333–348.
18. Karasek R, Baker D, Marxer F, Ahlbom A, Theorell T. Job decision latitude, job demands, and cardiovascular disease: a prospective study of Swedish men. *Am J Public Health.* 1981;71:694–705.
19. Karasek RA, Russell RS, Theorell T. Physiology of stress and regeneration of job related cardiovascular illness. *J Human Stress.* March 1982;29–42.
20. Karasek RA, Theorell T, Schwartz J, Pieper C, Alfredsson L. Job, psychological factors and coronary heart disease. *Adv Cardiol.* 1982;29:62–67.
21. Alfredsson L, Karasek R, Theorell T. Myocardial infarction risk and psychosocial environment: an analysis of the male Swedish working force. *Soc Sci Med.* 1982;16:463–467.
22. House JA, McMichael AJ, Wells JA, Kaplan BH, Landerman LR. Occupational stress and health among factory workers. *J Health Soc Behav.* 1979;20:139–160.
23. Cobb S, Rose RM. Hypertension, peptic ulcer, and diabetes in air traffic controllers. *JAMA.* 1973;224:489–492.
24. LaCroix AZ, Haynes SG. Gender differences in the stressfulness of workplace roles: a focus on work and health. In: Barnett R, Baruch G, Beiner L, eds. *Gender and Stress.* New York, NY: Free Press; 1987:96–121.
25. Schnall PL, Pieper C, Schwartz JE, et al. The relationship between "job strain," workplace diastolic blood pressure, and left ventricular mass index: results of a case-control study. *JAMA.* 1990;263:1929–1935.
26. Winkleby MA, Ragland DR, Syme SL. Self-reported stressors and hypertension: evidence of an inverse association. *Am J Epidemiol.* 1988;127:124–134.
27. Meyer E, Derogatis L, Miller M, Reading A. Hypertension and psychological stress. *Psychosomatics.* 1978;19:160–168.
28. Linden W, Feuerstein M. Essential hypertension and social coping behavior: experimental findings. *J Human Stress.* 1983;7:22–31.
29. House JS, Strecher V, Metzner HL, Robbins CA. Occupational stress and health among men and women in the Tecumseh community health study. *J Health Soc Behav.* 1986;27:62–77.
30. Schwartz JE, Pieper CF, Karasek RA. A procedure for linking psychosocial job characteristics data to health surveys. *Am J Public Health.* 1988;78:904–909.
31. Alfredsson L, Spetz C-L, Theorell T. Type of occupation and near-future hospitalization for myocardial infarction and some other diagnoses. *Int J Epidemiol.* 1985;14:378–388.
32. Pieper C, LaCroix AZ, Karasek RA. The relation of psychosocial dimensions of work with coronary heart disease risk factors: a meta analysis of five United States databases. *Am J Epidemiol.* 1989;129:483–494.
33. Winkleby MW, Ragland DR, Fisher JM, Syme SL. Excess risk of sickness and disease in bus drivers: a review and synthesis of epidemiologic studies. *Int J Epidemiol.* 1988;17:255–262.
34. Ragland DR, Winkleby M, Schwalbe J, et al. Prevalence of hypertension in bus drivers. *Int J Epidemiol.* 1987;16:1–7.
35. Casey BH, Poigner SB. *Measurement of Blood Pressure: A Manual for Training and Certification of Observers.* Houston, Tex:

Coordinating Center; 1976. Procedures of the Hypertension Detection and Follow-up Program.
36. Karasek RA, Schwartz J, Theorell T. *Job Characteristics, Occupation, and Coronary Heart Disease. Final Report—Phase 1.* New York, NY: Columbia University, Dept of Industrial Engineering and Operations Research; 1982.
37. Karasek RA, Gordon G, Peitrokovshy C, et al. *Job Content Instrument Questionnaire and User's Guide.* Revision 1.1. New York, NY: Columbia University, Job/Heart Project; March 1985.
38. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods.* New York, NY: Van Nostrand Reinhold Co; 1982.
39. Marmot MG, Rose G, Shipley M, Hamilton PJS. Employment grade and coronary heart disease in British civil servants. *J Epidemiol Community Health.* 1978;32:244–249.
40. Marmot MG. Psychosocial factors and blood pressure. *Prev Med.* 1985;14:451–465.
41. Syme SL, Oakes TW, Friedman GD. Social class and differences in blood pressure. *Am J Public Health.* 1974;64:619–620.
42. Lee RE, Schneider RF. Hypertension and arteriosclerosis in executive and nonexecutive personnel. *JAMA.* 1958;167:1447–1450.

APPENDIX A—Definitions of Covariables

Marital Status = married, separated, divorced, widowed, and single/never married.
Education = less than 11th grade, high school graduate, technical school/some college, college graduate or higher.
Income = rank-ordered into four categories, less than $25,000 to $35,000 or more.
Race = Black, White, Asian, Hispanic, and other.
Body Mass Index = Weight in kilograms, divided by height in meters squared.
Family History of Heart Disease = positive if participant's mother, father, or siblings ever had high blood pressure, heart disease, stroke, or a cardiovascular disease-related death.
Fitness = hours per week of moderately hard physical activity.
Alcohol Consumption = total drinks of wine, beer, or hard liquor in an average week.
Caffeine Intake = total cups of coffee or soft drinks per day.

APPENDIX B—Comparison of Means and Standard Deviations

Study	Job Demands	Decision Latitude
Karasek et al.[34]	34.5[a] ±8.5	72.9[a]±13.1
	30.5[a] ±8.5	71.3[a]±13.5
Reed et al.[15]	30.3 ±8.1	71.9 ± 8.1
Schnall et al.[35]	30.9 ±6.1	35.4 ± 6.1
Albright et al. (present study)	32.02±8.0	30.8 ± 5.0

[a]These figures are the computed averages of aggregated scores for different occupations.

Call for Abstracts for Epidemiology Late-Breaker Sessions

Exchange Session

The Epidemiology Section will sponsor a "late-breaker" Epidemiology exchange on Wednesday, November 11, 1992, during APHA's 1992 annual meeting in Washington, DC. The exchange will provide a forum for presentation of investigations or methods that have been conceived, conducted, and/or concluded so recently that members could not meet the deadline for abstract submission to other Epidemiology sessions. Abstracts should report on work conducted during the last 6–12 months.

Abstracts of fewer than 200 words (any format) and a stamped, self-addressed return envelope should be submitted to Polly A. Marchbanks, PhD, Chief, EIS Program, Epidemiology Program Office, Mailstop C08, Centers for Disease Control, Atlanta, GA 30333; (404) 639-3588.

Abstracts must be *received* by September 28, 1992. Decisions will be made by October 5, 1992.

Poster Session

The Epidemiology Section will again be sponsoring a late-breaker poster session at the annual meeting. The session will take place on Wednesday, November 11, 1992. Work completed in 1992 is eligible for consideration. Submit an abstract of fewer than 200 words (any format) and a *return envelope* to Cathey Falvo, MD, MPH, Graduate School of Health Science, Munger Pavilion, New York Medical College, Valhalla, NY 10595; tel. (914) 993-4323.

Abstracts must be *received* by October 2, 1992. Decisions will be mailed October 5, 1992. Students and recent graduates are particularly encouraged to submit abstracts.

Use of Hair Coloring Products and the Risk of Lymphoma, Multiple Myeloma, and Chronic Lymphocytic Leukemia

Shelia Hoar Zahm, ScD, Dennis D. Weisenburger, MD, Paula A. Babbitt, Robert C. Saal, Jimmie B. Vaught, PhD, and Aaron Blair, PhD

ABSTRACT

Objectives. Hair coloring products are widely used and contain components that are mutagenic and carcinogenic. An association between occupational exposure to hair coloring products and hematopoietic cancers has been reported, but the risk for these cancers among users has not been carefully evaluated.

Methods. We conducted a population-based, case-control study with telephone interviews from 385 non-Hodgkin's lymphoma cases, 70 Hodgkin's disease cases, 72 multiple myeloma cases, 56 chronic lymphocytic leukemia cases, and 1432 controls.

Results. Among women, use was associated with odds ratios of 1.5 for non-Hodgkin's lymphoma, 1.7 for Hodgkin's disease, 1.8 for multiple myeloma, and 1.0 for chronic lymphocytic leukemia. Risk was higher for permanent hair coloring products than for semi- or nonpermanent products, particularly for dark colors. Long duration and early age of first use tended to increase risk, but the patterns were inconsistent. Use was much less common in men and did not significantly increase risk.

Conclusions. The use of hair coloring products appears to increase the risk of non-Hodgkin's lymphoma. Multiple myeloma and Hodgkin's disease were also associated, although based on far fewer subjects. If these results represent a causal association, use of hair coloring products would account for 35% of non-Hodgkin's lymphoma cases in exposed women and 20% in all women. (*Am J Public Health.* 1992; 82:990–997)

Introduction

Hair coloring products are widely used and are known to contain components that are mutagenic and carcinogenic in animals.[1–4] Associations between occupational exposure to hair coloring products and hematopoietic cancer have been reported,[5–9] but little is known about the risks for hematopoietic cancers among users of such products. Two case-control studies have suggested that use of hair coloring products may increase the risk of leukemia and non-Hodgkin's lymphoma.[10,11] We evaluated risks associated with the use of hair coloring products in a population-based, case-control study of incident lymphoma, multiple myeloma, and chronic lymphocytic leukemia conducted in eastern Nebraska. Our report includes both sexes and involves more detailed data on personal use of hair coloring products than earlier studies.

Methods

Cases of non-Hodgkin's lymphoma, Hodgkin's disease, multiple myeloma, and chronic lymphocytic leukemia among White men and women, aged 21 years or older, residing in the 66 counties of eastern Nebraska, and diagnosed between July 1, 1983, and June 30, 1986, were identified through the Nebraska Lymphoma Study Group and area hospitals. Details concerning completeness of reporting and other study methods have been published elsewhere.[12] There were 426 non-Hodgkin's lymphoma, 73 Hodgkin's disease, 82 multiple myeloma, and 69 chronic lymphocytic leukemia cases histologically confirmed. Non-Hodgkin's lymphoma cases were reviewed and classified according to the Working Formulation[13] and underwent immunologic typing.[14,15] The various histologic types of non-Hodgkin's lymphoma were grouped as follicular (29%), diffuse (68%), or not otherwise specified (3%). The cases were also categorized histologically into follicular center cell (Working Formulation categories B through G and J: 71%) and nonfollicular center cell (categories A, H, I, and K: 29%) groups.[16] Composite lymphomas with follicular and diffuse components of the same cell type were assigned to the follicular category. Composite lymphomas with different cell types were assigned to the more indolent cell type. B-cell lymphomas accounted for 81% of the non-Hodgkin's lymphoma cases, whereas 8% of the cases were T-cell lymphomas. The remaining cases could not be classified: 5% were of indeterminant immunologic type and 5% had no specimens available for study.

The controls were selected from residents of the same 66-county area by 3:1 frequency matching by race, sex, vital status, and age (± 2 years) to the combined age distribution of the four cancer case series (non-Hodgkin's lymphoma, Hodg-

Shelia Hoar Zahm and Aaron Blair are with the Occupational Studies Section, National Cancer Institute, Rockville, Md. Dennis D. Weisenburger is with the Department of Pathology and Microbiology and the Eppley Institute for Research in Cancer and Allied Diseases, and Paula A. Babbitt was with the Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha. Robert C. Saal and Jimmie B. Vaught are with Westat Inc, Rockville, Md.

Requests for reprints should be sent to Shelia Hoar Zahm, ScD, Occupational Studies Section, National Cancer Institute, Executive Plaza North, Room 418, Rockville, MD 20892.

This paper was submitted to the Journal August 20, 1991, and accepted with revisions December 10, 1991.

TABLE 1—Number of Cases and Controls and Odds Ratios for Non-Hodgkin's Lymphoma, Hodgkin's Disease, Multiple Myeloma, and Chronic Lymphocytic Leukemia, by Sex and Hair Coloring Product Use

		Non-Hodgkin's Lymphoma		Hodgkin's Disease		Multiple Myeloma		Chronic Lymphocytic Leukemia	
	No. of Controls	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)
Women									
Never used any hair coloring product	373	74	1.0 ...	19	1.0 ...	14	1.0 ...	10	1.0 ...
Ever used any hair coloring product[b]	322	106	1.5 (1.1, 2.2)	16	1.7 (0.7, 4.0)	24	1.8 (0.9, 3.7)	9	1.0 (0.3, 2.6)
Semi- or nonpermanent hair coloring products	250	78	1.4 (1.0, 2.1)	8	1.2 (0.4, 3.2)	18	1.7 (0.8, 3.6)	7	0.9 (0.3, 2.8)
Permanent hair coloring products	109	41	1.7 (1.1, 2.8)	12	3.0 (1.1, 7.9)	11	2.8 (1.1, 7.1)	2	0.8 (0.1, 4.0)
Products that gradually change hair color	17	5	1.4 (0.4, 4.3)	0	... ...	2	3.1 (0.4, 17.0)	0	... ...
Men									
Never used any hair coloring product	675	190	1.0 ...	32	1.0 ...	27	1.0 ...	34	1.0 ...
Ever used any hair coloring product[b]	48	11	0.8 (0.4, 1.6)	3	1.7 (0.4, 6.3)	4	1.8 (0.5, 5.7)	3	1.0 (0.2, 3.8)
Semi- or nonpermanent hair coloring products	7	2	1.0 (0.1, 5.1)	0	... ...	1	3.6 (0.2, 33.8)	1	3.2 (0.1, 28.6)
Permanent hair coloring products	1	0	... ...	1	8.2 (0.2, 363.0)	1	29.1 (0.7, 1207.7)	0	... ...
Products that gradually change hair color	45	11	0.8 (0.4, 1.7)	2	1.3 (0.2, 5.9)	2	0.9 (0.1, 4.1)	2	0.7 (0.1, 3.2)

[a]Age adjusted.
[b]Subjects could have used more than one type of hair coloring product and therefore may contribute to more than one category.

kin's disease, multiple myeloma, and chronic lymphocytic leukemia). For living cases under age 65, 472 controls were selected by two-stage random-digit dialing.[17] For living cases aged 65 or older, 603 controls were selected from Health Care Financing Administration (Medicare) records. For deceased cases, 580 controls were selected from the Nebraska state mortality files using the additional matching factor of year of death. Persons with an underlying cause of death of non-Hodgkin's lymphoma, Hodgkin's disease, multiple myeloma, leukemia, malignancy of unknown site, aplastic anemia, suicide, homicide, or legal intervention were excluded as controls. A total of 1655 controls was selected.

Telephone interviews were conducted with 385 (201 men, 184 women) non-Hodgkin's lymphoma cases, 70 (35 men, 35 women) Hodgkin's disease cases, 72 (32 men, 40 women) multiple myeloma cases, 56 (37 men, 19 women) chronic lymphocytic leukemia cases, and 1432 (725 men, 707 women) controls, or their next of kin. The interview response rates were 90%, 96%, 88%, 81%, and 87%, respectively. The overall control response rate was 84% (a weighted average accounting for the 91% response rate in the household census phase of the random-digit dialing procedure, the 94% response rate of the randomly selected eligible random-digit dialing controls, the 84% response rate of the Health Care Financing Administration controls, and the 83% response rate of next of kin of deceased controls).

The interview questions included the ages of first and last use of products that gradually change hair color, of nonpermanent or semipermanent hair coloring products (defined as products that rinse out after a few shampoos), and of permanent hair coloring products (defined as products that last until the hair grows out). Users of non- or semipermanent products and permanent products were also asked the usual frequency of use and the basic color used most often. Because being diagnosed with cancer could affect use of hair coloring products, exposure histories for all cases and controls were truncated as of December 31, 1982, which precedes the earliest diagnosis date for all cases. Other known and suspected risk factors for the cancers under study, such as family history of cancer and pesticide exposures, were also included in the interview.

The measure of association was the odds ratio (OR). All risk estimates were adjusted for the effects of age by stratification (21 through 39, 40 through 59, 60 through 69, 70 through 79, and greater than 79 years). Maximum likelihood estimates of the overall risk and 95% confidence intervals (CIs) were computed by Gart's method.[18] For duration–response and frequency–response relationships, significance was assessed by means of Mantel's one-tailed linear trend test.[19] Logistic regression analyses were conducted using an SAS procedure for polychotomous outcomes, a method that accounts for the use of a single control group for the four case groups in its estimation of variance.[20] Attributable risks and CIs were calculated using the method proposed by Whittemore.[21]

Results

Table 1 presents the number of cases and controls and the ORs by sex and hair coloring product type. Among women, use of any hair coloring product was associated with non-Hodgkin's lymphoma, Hodgkin's disease, and multiple myeloma. Non-Hodgkin's lymphoma was significantly associated with the use of both semi- and nonpermanent hair coloring products (OR = 1.4) and permanent

TABLE 2—Number of Cases and Controls and Odds Ratios for Non-Hodgkin's Lymphoma, Hodgkin's Disease, Multiple Myeloma, and Chronic Lymphocytic Leukemia among Women by the Most Frequently Used Hair Coloring Product Color

		Non-Hodgkin's Lymphoma		Hodgkin's Disease		Multiple Myeloma		Chronic Lymphocytic Leukemia	
	No. of Controls	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)
Never used any hair coloring product	373	74		19		14		10	
Semi- or nonpermanent hair coloring product users[b]									
Blonde	41	10	1.1 (0.5, 2.5)	3	1.8 (0.4, 8.1)	2	1.3 (0.2, 6.8)	1	1.1 (0.05, 9.1)
Brown/brunette	138	44	1.4 (0.9, 2.2)	4	1.0 (0.3, 3.7)	9	1.6 (0.6, 4.1)	4	1.0 (0.3, 3.6)
Black	11	5	2.1 (0.6, 6.9)	0	... ...	2	3.9 (0.5, 22.1)	0	... ...
Red	6	4	2.5 (0.6, 10.5)	0	... ...	2	7.4 (0.9, 52.0)	0	... ...
Other	51	9	0.8 (0.3, 1.8)	1	1.7 (0.1, 17.6)	3	1.3 (0.3, 5.2)	2	1.1 (0.2, 5.9)
Unknown	3	6	... ...	0	... ...	0	... ...	0	... ...
Permanent hair coloring product users[b]									
Blonde	44	8	0.9 (0.4, 2.2)	6	2.5 (0.7, 8.5)	2	1.4 (0.2, 7.6)	0	... ...
Brown/brunette	43	20	2.0 (1.0, 3.8)	3	2.6 (0.5, 12.3)	6	3.8 (1.2, 11.6)	2	1.8 (0.3, 9.9)
Black	5	4	4.1 (0.9, 18.8)	0	... ...	0	... ...	0	... ...
Red	4	3	3.0 (0.5, 16.8)	0	... ...	1	5.4 (0.2, 68.5)	0	... ...
Other	12	6	2.4 (0.8, 7.5)	3	8.8 (1.1, 62.8)	2	4.1 (0.5, 23.9)	0	... ...
Unknown	1	0	... ...	0	... ...	0	... ...	0	... ...

[a]Age adjusted.
[b]Subjects could have used semi- or nonpermanent and permanent hair coloring products, and, therefore, may appear in the table twice.

hair coloring products (OR = 1.7). Women who used permanent hair coloring products only had an OR for non-Hodgkin's lymphoma of 1.9 (95% CI = 1.1, 3.4). Hodgkin's disease and multiple myeloma showed significant associations with the use of permanent hair coloring products. Women who had used permanent hair coloring products only had an OR of 3.4 (95% CI = 1.1, 10.7) for Hodgkin's disease, while users of semi- or nonpermanent products had an OR of 0.8. Chronic lymphocytic leukemia was not associated with the use of any hair coloring products. For all cancer types among women, except chronic lymphocytic leukemia, the risks associated with permanent hair coloring products were higher than those for semi- or nonpermanent hair coloring products, although CIs overlap.

Among men, the ORs for hair coloring product use were 0.8 for non-Hodgkin's lymphoma, 1.7 for Hodgkin's disease, 1.8 for multiple myeloma, and 1.0 for chronic lymphocytic leukemia. Non-Hodgkin's lymphoma was elevated among men who were long-term users of products that gradually changed hair color (1 year: OR = 1.0; 2 through 5 years: OR = 0.2; 6+ years: OR = 1.6; 95% CI = 0.5, 5.0), but neither their risk nor the duration trend (P = .362) was significant. The number of men using other hair coloring products was much less than the number of women and precluded meaningful analysis. The remaining results presented are for women only.

Risk varied by the most frequently used hair coloring product color (Table 2). For non-Hodgkin's lymphoma and multiple myeloma, the risks were higher among women who most often used brown or brunette, black, or red hair coloring products. In general, the risks associated with these darker color products were higher among users of permanent hair coloring products than semi- or nonpermanent products.

The risk of non-Hodgkin's lymphoma increased significantly with the duration of use and younger age at first use of semi- or nonpermanent products, although the risks dropped off in the highest category of both variables (Table 3). Risk also increased significantly with duration of use of dark permanent hair coloring products and younger age at first use of permanent hair coloring products, but did not show consistent exposure–response gradients (Table 4). Risk did not increase with frequency of use. The same patterns, with higher levels of risk, were seen when the analyses were restricted to black, brown/brunette, and red hair coloring products only. In general, users of permanent hair coloring products had a higher risk of non-Hodgkin's lymphoma relative to women who never used any hair coloring products and relative to women who used semi- or nonpermanent hair coloring products at most levels of duration, age, and frequency of use. Women who began using hair coloring products prior to 1970 had slightly lower risks (semi- or nonpermanent: OR = 1.3; permanent: OR = 1.6) than women who began using hair coloring products after that date (semi- or nonpermanent: OR = 1.5; permanent: OR = 2.2).

Because many of the age and time characteristics were interrelated, we performed polychotomous logistic regression entering all the variables presented in Tables 1 through 4. The results of the regression analyses were similar to the analyses already presented. Color of hair coloring products used, duration of semi- or nonpermanent hair coloring product use, decreasing age of first use of permanent hair coloring products, and age at diagnosis significantly increased the risk of non-Hodgkin's lymphoma. The other age, duration, and frequency variables were not significantly associated with non-Hodgkin's lymphoma.

Women who had ever used semi- or nonpermanent hair coloring products had a higher risk of follicular center cell lymphoma (OR = 1.6; 95% CI = 1.0, 2.5) than those who developed nonfollicular center cell lymphoma (OR = 1.1; 95% CI = 0.6, 2.0). Users of permanent hair coloring

TABLE 3—Number of Cases and Controls and Odds Ratios for Non-Hodgkin's Lymphoma, Hodgkin's Disease, and Multiple Myeloma among Women, by Characteristics of Semipermanent or Nonpermanent Hair Coloring Product Use

	Non-Hodgkin's Lymphoma						Hodgkin's Disease				Multiple Myeloma			
	All Colors			Black, Brown/Brunette, Red Only			All Colors		Black, Brown/Brunette, Red		All Colors		Black, Brown/Brunette, Red	
	No.	No. of Controls	Odds Ratio[a] (95% CI)	No.	No. of Controls	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)
Never used any hair coloring product	74	373	1.0 ...	74	373	1.0 ...	19	1.0 ...	19	1.0 ...	14	1.0 ...	14	1.0 ...
Duration of use, y														
1–10	29	113	1.1 (0.7, 1.9)	19	75	1.1 (0.6, 2.0)	5	1.1 (0.3, 3.6)	2	0.7 (0.1, 3.7)	10	2.2 (0.9, 5.8)	6	2.1 (0.7, 6.4)
11–20	27	57	2.0 (1.2, 3.6)	18	36	2.1 (1.1, 4.2)	1	0.8 (0.04, 7.1)	1	1.5 (0.1, 13.9)	3	1.3 (0.3, 4.9)	3	2.0 (0.4, 8.0)
21+	11	40	1.2 (0.6, 2.7)	9	19	2.2 (0.9, 5.6)	2	5.0 (0.5, 40.2)	1	4.6 (0.2, 54.0)	2	1.0 (0.1, 5.0)	1	1.1 (0.1, 8.8)
Unknown	9	32	... ...	6	20	... ...	0	... ...	0	... ...	2	... ...	2	... ...
P value for trend			.024			.004		.154		.239		.320		.174
Age at first use, y														
61+	9	46	0.9 (0.4, 2.0)	4	31	0.6 (0.2, 1.8)	1	1.6 (0.1, 18.0)	1	2.4 (0.1, 27.4)	1	0.4 (0.02, 3.4)	1	0.7 (0.03, 5.2)
51–60	18	50	1.4 (0.7, 2.7)	15	31	1.9 (0.9, 4.0)	1	1.6 (0.1, 15.6)	0	... ...	4	1.8 (0.5, 6.2)	3	2.2 (0.5, 9.1)
41–50	20	52	1.6 (0.9, 3.0)	15	31	2.1 (1.0, 4.3)	0	... ...	0	... ...	5	2.2 (0.7, 7.3)	3	2.4 (0.5, 10.0)
≤40	20	67	1.4 (0.7, 2.6)	12	39	1.5 (0.7, 3.2)	6	1.5 (0.5, 4.7)	3	1.3 (0.3, 5.8)	6	3.1 (1.0, 9.5)	4	3.9 (1.0, 14.6)
Unknown	11	35	... ...	7	23	... ...	0	... ...	0	... ...	2	... ...	2	... ...
P value for trend			.030			.012		.300		.475		.010		.006
Frequency of use per year														
1–11	33	80	1.9 (1.1, 3.2)	25	54	2.1 (1.2, 3.7)	5	1.6 (0.4, 5.6)	2	1.1 (0.2, 6.3)	8	2.8 (1.0, 7.7)	6	3.1 (1.0, 9.5)
12–23	16	76	1.0 (0.5, 1.8)	11	50	1.0 (0.5, 2.2)	2	1.0 (0.1, 5.4)	1	0.7 (0.03, 6.2)	3	0.9 (0.2, 3.5)	2	0.9 (0.1, 4.6)
24–51	7	26	1.2 (0.4, 3.0)	4	16	1.2 (0.3, 4.0)	0	... ...	0	... ...	2	1.6 (0.2, 8.2)	2	2.9 (0.4, 15.4)
52+	16	48	1.4 (0.7, 2.7)	10	22	1.8 (0.7, 4.2)	1	1.5 (0.1, 16.9)	1	3.4 (0.1, 41.4)	2	0.9 (0.1, 4.4)	1	0.9 (0.04, 7.6)
Unknown	6	18	... ...	3	13	... ...	0	... ...	0	... ...	3	... ...	2	... ...
P value for trend			.177			.083		.493		.429		.497		.268

[a]Age adjusted.

TABLE 4—Number of Cases and Controls and Odds Ratios for Non-Hodgkin's Lymphoma, Hodgkin's Disease, and Multiple Myeloma among Women, by Characteristics of Permanent Hair Coloring Product Use

	Non-Hodgkin's Lymphoma						Hodgkin's Disease				Multiple Myeloma			
	All Colors			Black, Brown/Brunette, Red Only			All Colors		Black, Brown/Brunette, Red		All Colors		Black, Brown/Brunette, Red	
	No.	No. of Controls	Odds Ratio[a] (95% CI)	No.	No. of Controls	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)	No.	Odds Ratio[a] (95% CI)
Never used any hair coloring product	74	373	1.0 ...	74	373	1.0 ...	19	1.0 ...	19	1.0 ...	14	1.0 ...	14	1.0 ...
Duration of use, y														
1–10	23	45	2.4 (1.3, 4.6)	15	24	2.8 (1.3, 6.0)	9	3.6 (1.2, 11.1)	2	2.5 (0.3, 15.7)	3	1.9 (0.4, 8.2)	2	2.2 (0.3, 11.7)
11–20	11	24	2.2 (0.9, 5.0)	7	14	2.2 (0.8, 6.2)	2	2.8 (0.4, 17.9)	1	3.6 (0.1, 36.8)	5	5.6 (1.5, 19.6)	3	5.4 (1.1, 24.4)
21+	4	24	0.7 (0.2, 2.3)	4	7	2.3 (0.6, 9.4)	1	2.0 (0.1, 20.5)	0	... ...	2	2.3 (0.3, 12.1)	1	3.4 (0.1, 33.1)
Unknown	3	14	... ...	1	6	... ...	0	... ...	0	... ...	1	... ...	1	... ...
P value for trend			.102			.003		.018		.145		.008		.005
Age at first use, y[b]														
51+	9	28	1.4 (0.6, 3.3)	9	20	1.9 (0.7, 4.8)	0	... ...	0	... ...	2	1.7 (0.2, 8.8)	2	2.3 (0.3, 12.1)
41–50	11	14	3.3 (1.3, 8.5)	8	7	5.0 (1.5, 16.9)	0	... ...	0	... ...	7	13.4 (3.7, 49.5)	4	18.0 (2.3, 99.3)
31–40	10	31	1.4 (0.6, 3.2)	6	15	1.8 (0.6, 5.2)	3	4.1 (0.7, 22.4)	2	5.9 (0.7, 39.4)	0	... ...	0	... ...
≤30	8	23	2.1 (0.7, 5.6)	3	4	3.7 (0.6, 23.3)	9	4.1 (1.3, 12.9)	1	2.3 (0.1, 37.6)	1	2.0 (0.1, 23.6)	0	... ...
Unknown	3	13	... ...	1	6	... ...	0	... ...	0	... ...	1	... ...	1	... ...
P value for trend			.013			.001		.001		.045		.042		.040
Frequency of use per year														
1–3	20	28	3.6 (1.8, 7.4)	10	10	4.6 (1.7, 12.6)	4	3.5 (0.8, 15.0)	2	10.5 (1.2, 76.3)	1	1.0 (0.05, 8.8)	0	... ...
4–6	9	20	2.0 (0.8, 5.0)	6	9	2.9 (0.9, 9.6)	7	7.5 (2.0, 29.1)	1	1.8 (0.1, 23.9)	3	4.2 (0.8, 19.4)	2	6.5 (0.8, 44.3)
7+	8	48	0.7 (0.3, 1.7)	7	27	1.1 (0.4, 2.8)	1	0.6 (0.03, 5.7)	0	... ...	4	2.4 (0.6, 9.0)	4	3.7 (0.9, 14.1)
Unknown	4	13	... ...	4	6	... ...	0	... ...	0	... ...	3	... ...	1	... ...
P value for trend			.339			.107		.052		.423		.030		.004

[a]Age adjusted.
[b]The age categories for first use of permanent hair coloring products differ from the categories used for semi- or nonpermanent hair coloring products because use patterns differ. For both analyses, the categories were selected to achieve as equal a distribution of the control subjects as possible.

products also had a higher risk of follicular center cell lymphoma (OR = 2.0; 95% CI = 1.1, 3.5) than those who developed nonfollicular center cell lymphoma (OR = 1.3; 95% CI = 0.6, 2.8). Women using dark permanent products also had the highest risk for follicular center cell lymphoma (OR = 2.5; 95% CI = 1.3, 4.9) and had an almost threefold increased risk (OR = 2.8; 95% CI = 1.2, 6.5) of developing large cell lymphoma of follicular center cell origin (categories D, G, and J).

Users of semi- or nonpermanent hair coloring products had similar risks for B-cell lymphoma (OR = 1.5; 95% CI = 1.0, 2.2) and for T-cell lymphoma (OR = 1.5; 95% CI = 0.3, 7.7), whereas users of permanent hair coloring products had a higher risk for T-cell lymphoma (OR = 3.5; 95% CI = 0.6, 19.3) than for B-cell lymphoma (OR = 1.7; 95% CI = 1.0, 2.9). However, the ORs were not significantly different from each other.

Adjustments for family history of cancer, cigarette smoking, herbicide use, and insecticide use did not change the risk estimates for non-Hodgkin's lymphoma and hair coloring product use.

Risk estimates for non-Hodgkin's lymphoma based on data from self-respondents or proxy respondents were similar for use of any hair coloring products (self: OR = 1.5; proxy: OR = 1.6), semi- or nonpermanent products (self: OR = 1.4; proxy: OR = 1.4), permanent products (self: OR = 1.7; proxy: OR = 1.9), and products that changed hair color gradually (men, self: OR = 0.8; proxy: OR = 0.9). Risks and trends associated with age at first use and duration of semi- or nonpermanent hair coloring products and age at first use of permanent products were slightly stronger among self-respondents than among subjects whose data were supplied by proxies. No meaningful differences were seen between self-respondents and proxies for frequency and color of semi- or nonpermanent hair coloring products. Conversely, risks by duration, frequency, and color of permanent hair coloring products were slightly higher among subjects whose data were supplied by proxies than among self-respondents.

There were far fewer cases of the other three cancers than of non-Hodgkin's lymphoma, which limited our ability to investigate risks by use characteristics. Nevertheless, significant trends in risk for multiple myeloma were observed for decreasing age at first use of semi- or nonpermanent hair coloring products (Table 3) and duration, decreasing age at first use, and frequency of use of permanent hair coloring products (Table 4); however, at least in part because of small numbers, most risk patterns were inconsistent. Polychotomous logistic regression analyses, based on all variables in Tables 1 through 4, revealed only color to be significantly associated with risk of multiple myeloma. Nonsignificant associations were observed with decreasing age at first use of semi- or nonpermanent-or permanent hair coloring products and frequency of use of permanent hair coloring products.

A significant trend was seen for Hodgkin's disease and decreasing age at first use of permanent hair coloring products (Table 4), but the significant trend was primarily due to the lack of Hodgkin's disease cases in the older age at first use categories. The Hodgkin's disease cases were younger on average than the controls, who were frequency matched to the combined group of four cancer types; thus, the difference observed by age at first use may not have etiologic significance. The logistic regression analyses showed age at diagnosis and decreasing age at first use of permanent hair coloring products to be significantly associated with development of Hodgkin's disease. Nonsignificant associations were seen with duration and younger age at first use of semi- or nonpermanent hair coloring products.

Discussion

Experimental studies have shown that hair coloring products contain mutagenic and carcinogenic compounds, which vary by hair coloring product type and color.[1-4] Although some of these components have been removed from use over the last two decades, many carcinogenic aromatic, nitroso, and amino compounds remain and are found in both semipermanent and permanent hair coloring products. Permanent hair coloring products also contain oxidizing agents that allow the dye to irreversibly bind to the hair shaft. This generally lowers absorption by the skin of specific aromatic amines, but the oxidation process also results in the formation of some products, such as the three-ringed "Bandrowski-base," that are absorbed.[1] Many different chemical dyes can be found in most color formulations, but they usually occur in greater concentrations in the black, brown, and red dyes than in blond and silver bleach toner formulations. It is interesting that, in the present epidemiologic study, the variation in cancer risks by product type and color is somewhat consistent with what would be expected based on the concentration of the constituents and biochemical action of the product formulations. Hair coloring products that gradually change hair color have been shown to be absorbed by the users.[22] For example, gradual products contain lead acetate[23,24] and result in absorption of lead by the users.[22] Although the target sites for carcinogenic hair coloring components in feeding studies of rats and mice have generally been the urinary tract, liver, mammary gland, skin, forestomach, and thyroid gland,[25] there have been reports of increased lymphoma in mice in two skin painting studies of hair dyes.[26,27]

In the present study, significant risks were observed for overall use of hair coloring products, darker colors, and trends of some use characteristics, but the patterns of risk by some use characteristics were inconsistent. These inconsistencies could arise because the association is noncausal, because of chance (numbers of exposed cases were small), or because of misclassification of self- and proxy-reported exposure data for some of the use characteristics. There is little information in the scientific literature on the quality of self- or proxy-reported histories of hair coloring product use. In a small reliability study of 32 women (16 cases and 16 controls), the correlation between reported duration of hair dye use from two interviews 1 year apart was .86 and was similar for cases and controls.[28] The correlation for frequency of use (.92) was slightly higher. We have no such data on our subjects, and no information on recall over longer time periods or by surrogate respondents. It seems reasonable, however, to assume that the quality of reporting on type and color of hair coloring product used would be more accurate than reporting on age of first use or annual frequency of use. Recall worsens with the degree of detail required.[29] The type of misclassification of the use characteristics likely to occur would tend to reduce ORs and dampen or disrupt trends.[30,31] In the present study, there was little difference between self-respondents and proxy respondents for risk estimates of non-Hodgkin's lymphoma for the overall "ever" use of hair coloring products. Although the number of subjects in each category was small and risk estimates were highly variable, we did examine use characteristics by respondent type. Slightly greater risks and trends were seen among self-respondents than among proxies for the age at first use of any hair coloring product and duration of semipermanent hair coloring products.

These data suggest that inclusion of proxies, with their potentially less accurate data, may have resulted in underestimates of true risks and dampened trends for the detailed use characteristics. However, proxies yielded slightly higher risk estimates for duration, frequency, and color of permanent hair coloring products than self-respondents. Similar results for subjects and proxies for semipermanent frequency variables suggest that the data from the two groups may be equally poor.

Case response bias needs to be considered as a possible explanation for the associations observed in this study. With several cancers showing an association, one cannot easily dismiss case response bias completely. Shore et al.,[28] however, found little evidence for case response bias for hair dye use in breast cancer patients in 1979. More recent data on potential for case response bias with respect to hair coloring product use are not available. Awareness of potential health risks associated with hair coloring products may have increased. It seems unlikely, however, that case response bias would operate in a manner so as to yield risks by type and color of product that are consistent with what might be expected based on the concentration of oxidizers and dyes in the formulations, facts not likely to be known by study subjects.

We were able to adjust for the influence of several known or suspected risk factors for non-Hodgkin's lymphoma and the other cancers and found no evidence of confounding. It has been hypothesized that the hair graying or other conditions that lead to the use of hair coloring products are a reflection of hormonal status or other factors that themselves are responsible for increased cancer risk.[32] While this may be relevant for breast cancer, there is no evidence that hormonal factors play a role in the development of non-Hodgkin's lymphoma,[33] Hodgkin's disease,[34] or multiple myeloma.[35] A single report links the use of hair dyes to connective tissue disorders, such as systemic lupus erythematosus and scleroderma.[36] These acquired disorders of immunity have also been associated with an increased risk of non-Hodgkin's lymphoma[33] and multiple myeloma, although the association with multiple myeloma was observed in clinical series[37] but not case-control studies.[38–40] These findings suggest that hair coloring products may act by affecting the immune system. Freni-Titulaer et al.[36] hypothesized that the association between connective tissue disease and hair coloring products may be affected by acetylator phenotype. Slow acetylators would be at greater risk than fast acetylators. The aromatic amines in hair coloring products, which are absorbed through the scalp, are metabolized through acetylation. Genetically determined metabolic phenotypes could affect risk, dampening apparent associations in populations that are heterogeneous for the metabolic phenotype.

Most epidemiologic studies of hair coloring products and hematopoietic cancers have focused on people who were occupationally exposed (i.e., cosmetologists). The relative exposure levels of professional compared with personal use are unknown. If cosmetologists use protective gloves, their dermal exposures may, in fact, be less than those of persons whose hair is being colored. Two reports of significant excesses of multiple myeloma,[7,9] one report of a significant excess of leukemia,[9] and several reports of nonsignificant excesses of all hematopoietic cancers combined[5,8,41–43] have been reported among cosmetologists.

Far fewer data are available on the risk of hematopoietic cancers among users of hair coloring products. A study of registered nurses had fewer than expected lymphoma cases (10 vs 15.7) among women who had used permanent hair dyes.[44] Another study, which combined lymphoma and leukemia, reported an excess risk in London but a deficit risk in Toronto associated with use of permanent or semipermanent dyes.[45] The study by Cantor et al.[11] of non-Hodgkin's lymphoma and leukemia is by far the largest case-control study to address the risk for these tumors in hair coloring product users. Both cancer types were associated with hair dye use, but the study was restricted to men. The present study included only 21 male cases (11 non-Hodgkin's lymphoma, 3 Hodgkin's disease, 4 multiple myeloma, and 3 chronic lymphocytic leukemia) who reported use of hair coloring products; therefore, we were unable to thoroughly evaluate their risk or the consistency of the results for men between the two studies. Our data suggest that men are more likely to use products that gradually change hair color (6% of controls) than other types of hair coloring products (1% of controls). More research is needed on the risk of these gradual hair coloring products. There were only 12 cases of chronic lymphocytic leukemia (9 women, 3 men) exposed to hair coloring products in the present study, too few to assess risk with any confidence. The study of Cantor et al.[11] found at least a twofold excess risk of leukemia associated with hair coloring product use in men for a number of different leukemia subtypes, but not chronic lymphocytic leukemia and acute nonlymphocytic leukemia. A case-control study of acute nonlymphocytic leukemia conducted in Baltimore, however, found an association with use of hair dyes among both men and women.[10] To our knowledge, there are no reports on the risk of multiple myeloma associated with the use of hair coloring products, although there are several reports of excesses of multiple myeloma in persons occupationally exposed to hair dyes.[35]

With the widespread use of hair coloring products among women (46% of the controls), the question of cancer risk is an important public health issue. Using the Whittemore method[21] for calculating attributable risks and CIs, based on a relative risk of 1.53 for non-Hodgkin's lymphoma for ever having used any hair coloring product, these exposures would account for 35% (95% CI = 9%, 52%) of the non-Hodgkin's lymphoma occurring in exposed women and 20% (95% CI = 3%, 35%) in all women. Non-Hodgkin's lymphoma is one of the most rapidly increasing cancers in developed countries,[46,47] second in the United States only to melanoma and female lung cancer.[48] Between 1973 and 1987, the incidence of non-Hodgkin's lymphoma in the United States increased by more than 50%.[48] Some, but not the majority, of this increase is due to lymphomas arising in patients with the acquired immunodeficiency syndrome. Historical data on trends in hair coloring product use, which would help assess its role in the changing occurrence of non-Hodgkin's lymphoma, are not available. In the United States, however, current estimates of the percentage of women who have used or are using hair coloring products range from 20% (G. N. McEwen, Jr, personal communication, February 1991) to 40%.[49] The excess of non-Hodgkin's lymphoma in the present study is consistent with findings from the largest case-control study of this tumor and non-occupational hair dye exposure.[11] If it is causal, hair coloring product use would explain a larger percentage of non-Hodgkin's lymphoma cases among women than any other currently known or suspected risk factor. □

References

1. International Agency for Research on Cancer. *Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Man:*

Some Aromatic Amines and Related Nitro Compounds—Hair Dyes, Colouring Agents and Miscellaneous Industrial Chemicals. Lyon, France: International Agency for Research on Cancer; 1978; 16.

2. International Agency for Research on Cancer. *Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Man: Some Aromatic Amines, Anthraquinones and Nitroso Compounds, and Inorganic Fluorides Used in Drinking Water and Dental Preparations.* Lyon, France: International Agency for Research on Cancer; 1982; 27.
3. Kinlen L. Hair dyes: epidemiologic evidence. In: Wald NJ, Doll R, eds. *Interpretation of Negative Epidemiological Evidence for Carcinogenicity.* Lyon, France: International Agency for Research on Cancer; 1985. Scientific Publication Number 65.
4. International Agency for Research on Cancer. *Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Man: Overall Evaluation of Carcinogenicity: An Update of IARC Monographs, Volumes 1–42.* Lyon, France: International Agency for Research on Cancer; 1987.
5. Decoufle P, Stanislawczyk K, Houten L, Bross IDJ, Viadana E. *A Retrospective Survey of Cancer in Relation to Occupations.* Cincinnati, Ohio: US Government Printing Office; 1977. DHEW publication no. (NIOSH) 77–178.
6. Agu VU, Christensen BL, Buffler PA. Geographic patterns of multiple myeloma: racial and industrial correlates, state of Texas, 1969–71. *JNCI.* 1980;65:735–738.
7. Guidotti S, Wright WE, Peters JM. Multiple myeloma in cosmetologists. *Am J Ind Med.* 1982;3:169–171.
8. Teta MJ, Walrath J, Meigs JM, Flannery JT. Cancer incidence among cosmetologists. *JNCI.* 1984;72:1051–1057.
9. Spinelli JJ, Gallagher RP, Band PR, Threlfall WJ. Multiple myeloma, leukemia, and cancer of the ovary in cosmetologists and hair dressers. *Am J Ind Med.* 1984;6:97–102.
10. Markowitz JA, Szklo M, Sensenbrenner LL, Warm S. Hair dyes and acute nonlymphocytic leukemia (ANLL). *Am J Epidemiol.* 1985;122:523. Abstract.
11. Cantor KP, Blair A, Everett G, et al. Hair dye use and risk of leukemia and lymphoma. *Am J Public Health.* 1988;78:570–571.
12. Zahm SH, Weisenburger DD, Babbitt PA, et al. A case-control study of non-Hodgkin's lymphoma and the herbicide 2,4-dichlorophenoxyacetic acid (2,4-D) in eastern Nebraska. *Epidemiology.* 1990;1: 349–356.
13. The Non-Hodgkin's Lymphoma Pathologic Classification Project. National Cancer Institute sponsored study of classifications of non-Hodgkin's lymphomas: summary and description of a working formulation for clinical usage. *Cancer.* 1982; 49:2112–2135.
14. Linder J, Ye Y, Harrington DS, Armitage JO, Weisenburger DD. Monoclonal antibodies marking T lymphocytes in paraffin-embedded tissue. *Am J Pathol.* 1987;127: 1–8.
15. Linder J, Ye Y, Armitage JO, Weisenburger DD. Monoclonal antibodies marking B-cell non-Hodgkin's lymphoma in paraffin-embedded tissue. *Modern Pathol.* 1988;1:29–34.
16. Weisenburger DD, Harrington DS, Armitage JO. B-cell neoplasia: a conceptual understanding based on the normal humoral immune response. *Pathology Annu.* 1990; 25:99–115.
17. Waksberg J. Sampling methods for random digit dialing. *J Am Stat Assoc.* 1978;73:40–46.
18. Gart JJ. Point and interval estimation of the common odds ratio in the combination of 2×2 tables with fixed marginals. *Biometrika.* 1970;57:471–475.
19. Mantel N. Chi square test with 1 degree of freedom, extension of the Mantel-Haenszel procedure. *Am Stat Assoc J.* 1963;58: 690–700.
20. Jones RJ. Probability estimation using a multinomial logistic function. *J Stat Comput Simulation.* 1975;3:315–329.
21. Whittemore AS. Estimating attributable risk from case-control studies. *Am J Epidemiol.* 1983;117:76–85.
22. Marzulli FN, Watlington PM, Maibach HI. Exploratory skin penetration findings relating to the use of lead acetate hair dyes: hair as a test tissue for monitoring uptake of systemic lead. *Curr Probl Dermatol.* 1978; 7:196–204.
23. Searle CE, Harnden DG. Lead in hair-dye preparations. *Lancet.* 1979;ii(8151):1070.
24. Waldron HA. Lead poisoning from cosmetics. *Lancet.* 1979;ii(8151):1070–1071.
25. Sontag JM. Carcinogenicity of substituted-benzenediamines (phenylenediamines) in rats and mice. *JNCI.* 1981;66:591–602.
26. Jacobs MM, Burnett CM, Penicnak AJ, et al. Evaluation of the toxicity and carcinogenicity of hair dyes in Swiss mice. *Drug Chem Toxicol.* 1984;7:573–586.
27. Searle CE, Jones EL. Effects of repeated applications of two semipermanent hair dyes to the skin of A and DBAf mice. *Br J Cancer.* 1977;36:467–478.
28. Shore RE, Pasternack BS, Thiessen EU, Sadow M, Forbes R, Albert RE. A case-control study of hair dye use and breast cancer. *JNCI.* 1979;62:277–283.
29. Coughlin SS. Recall bias in epidemiologic studies. *J Clin Epidemiol.* 1990;43:87–91.
30. Marshall JR, Priore R, Graham S, Brasure J. On the distortion of risk estimates in multiple exposure level case-control studies. *Am J Epidemiol.* 1981;113:464–480.
31. Dosemeci M, Wacholder S, Lubin JH. Does nondifferential misclassification of exposure always bias a true effect toward the null value? *Am J Epidemiol.* 1990;132: 746–748.
32. Davies JM. Hair dyes and cancer: another confounding factor? *Lancet.* 1979;ii(8141): 536.
33. Greene MH. Non-Hodgkin's lymphoma and mycosis fungoides. In: Schottenfeld D, Fraumeni JF Jr, eds. *Cancer Epidemiology and Prevention.* Philadelphia, Pa: WB Saunders Co; 1982:754–778.
34. Grufferman S. Hodgkin's disease. In: Schottenfeld D, Fraumeni JF Jr, eds. *Cancer Epidemiology and Prevention.* Philadelphia, Pa: WB Saunders Co; 1982:739–753.
35. Riedel DA, Pottern LM, Blattner WA. The epidemiology of multiple myeloma. In: Wiernik PH, Canellos GP, Kyle RA, Schiffer CA, eds. *Neoplastic Diseases of the Blood.* 2nd ed. New York, NY: Churchill Livingstone Inc; 1991:347–372.
36. Freni-Titulaer LWJ, Kelley DB, Grow AG, McKinley TW, Arnett FC, Hochberg MC. Connective tissue disease in southeastern Georgia: a case-control study of etiologic factors. *Am J Epidemiol.* 1989;2: 404–409.
37. Blattner WA. Epidemiology of multiple myeloma and related plasma cell disorders: an analytic review. In: Potter M, ed. *Progress in Myeloma: Biology of Myeloma.* New York, NY: Elsevier/North Holland Inc; 1980:1–65.
38. Linet MS, Harlow SD, McLaughlin JK. A case-control study of multiple myeloma in Whites: chronic antigenic stimulation, occupation, and drug use. *Cancer Res.* 1987; 47:2978–2981.
39. Cuzick J, DeStavola BL. Autoimmune disorders and multiple myeloma. *Int J Epidemiol.* 1989;18:283.
40. Gallagher RP, Spinelli JJ, Elwood JM, Skipper DH. Allergies and agricultural exposure and risk factors for multiple myeloma. *Br J Cancer.* 1983;48:853–857.
41. Menck HR, Pike MC, Henderson BE, Jing JS. Lung cancer risk among beauticians and other female workers. *JNCI.* 1977;59: 1423–1425.
42. Kono S, Tokudome S, Ikeda M, Yoshimura T, Kuratsune M. Cancer and other causes of death among female beauticians. *JNCI.* 1983;70:443–446.
43. Guberan E, Raymond L, Sweetnam PM. Increased risk for male bladder cancer among a cohort of male and female hairdressers from Geneva. *Int J Epidemiol.* 1985;14:549–554.
44. Hennekens CH, Speizer FE, Rosner B, Bain CJ, Belanger C, Peto R. Use of permanent hair dyes and cancer among registered nurses. *Lancet.* 1979;i:1390–1393.
45. Stavraky KM, Clarke EA, Donner A. A case-control study of hair-dye use and cancers of various sites. *Br J Cancer.* 1981;43: 236–239.
46. Doll R. Are we winning the fight against cancer? An epidemiological assessment. *Eur J Cancer.* 1990;26:500–508.
47. Davis DL, Hoel D, Fox J, Lopez A. International trends in cancer mortality in France, West Germany, Italy, Japan, England and Wales, and the USA. *Lancet.* 1990;336:474–481.
48. Ries LAG, Hankey BF, Edwards BK. *Cancer Statistics Review, 1973–1987.* Washington, DC: US Government Printing Office; 1990. DHHS publication no. (NIH) 90-2789.
49. Ponte L. Why our hair turns gray. *Reader's Digest.* March 1991:87–89.

Erratum

In: Catania JA, Coates TJ, Kegeles S, et al. Condom use in multi-ethnic neighborhoods of San Francisco: the population-based AMEN (AIDS in Multi-Ethnic Neighborhoods) study. *Am J Public Health* 1992;82:284–287.

Because of programming errors detected by the authors only after publication of the article, Tables 4, 5, and 6 provided some incorrect data. The corrected data are shown in the tables printed here.

The substantive conclusions based on the original tables have not changed with the exception that the "Other" condom-related beliefs measure is no longer a significant correlate of condom use for gay and bisexual men. Both the discussion and the abstract are correct as published in the original article. The authors regret the errors.

TABLE 5—Psychosocial Correlates of Condom Use (Always vs Sometimes/Never)[a] for Heterosexual Women

	Odds Ratio	Confidence Interval 95%	*P* value[b]
Communication	1.9	1.4, 2.5	.001
Enjoyment	1.6	1.2, 2.2	.001
Condom attitudes	1.1	0.8, 1.3	ns
Self-efficacy beliefs	1.1	0.7, 1.4	ns
Religiosity	1.3	1.0, 1.8	ns
Tested for HIV	1.1	0.5, 2.4	ns
Monogamous	5.3	2.1, 13.5	.001
Susceptibility	0.99	0.7, 1.3	ns
AIDS anxiety	0.86	0.53, 1.3	ns
Black	0.24	0.11, 0.54	.001
Hispanic	0.25	0.09, 0.71	.01

[a]For those who always used condoms, n = 52; for those who sometimes or never used condoms, n = 528; Analyses based on multiple logistic regression.
[b]*P* value of the improvement chi-square at each step; ns = not significant.

TABLE 4—Psychosocial Correlates of Condom Use (Always vs Sometimes/Never)[a] for Heterosexual Men

	Odds Ratio	Confidence Interval 95%	*P* value[b]
Communication	1.8	1.4, 2.3	.001
Enjoyment	1.9	1.3, 2.5	.001
Condom attitudes	1.4	1.1, 1.7	.01
Self-efficacy beliefs	1.0	0.8, 1.3	ns
Religiosity	1.1	0.8, 1.4	ns
Tested for HIV	1.1	0.5, 2.4	ns
Monogamous	3.3	1.4, 7.7	.01
Susceptibility	1.1	0.90, 1.4	ns
AIDS anxiety	0.97	0.62, 1.4	ns
Black	1.4	0.61, 3.6	ns
Hispanic	2.1	0.91, 5.1	ns

[a]For those who always used condoms, n = 49; for those who sometimes or never used condoms, n = 495. Analyses were based on multiple logistic regression.
[b]*P* value of the improvement chi-square at each step; ns = not significant.

TABLE 6—Psychosocial Correlates of Condom Use (Always vs Sometimes/Never)[a] for Gay/Bisexual Men

	Odds Ratio	Confidence Interval 95%	*P* value[b]
Communication	1.5	1.1, 2.1	.05
Enjoyment	1.5	0.95, 1.9	.05
Condom attitudes	1.1	0.71, 1.6	ns
Self-efficacy beliefs	2.0	1.1, 3.6	.05
Religiosity	1.3	0.90, 2.1	ns
Tested for HIV	0.66	0.26, 1.7	ns
Monogamous	1.4	0.38, 4.9	ns
Susceptibility	1.2	0.77, 1.8	ns
AIDS anxiety	0.73	0.38, 1.4	ns
Black	0.45	0.10, 2.1	ns
Hispanic	0.50	0.11, 2.5	ns

[a]For those who always used condoms, n = 50; for those who sometimes or never used condoms, n = 55; Analyses based on multiple logistic regression.
[b]*P* value of the improvement chi-square at each step; ns = not significant.

Education, Race, and High-Density Lipoprotein Cholesterol among US Adults

David S. Freedman, PhD, David S. Strogatz, PhD, David F. Williamson, PhD, and Ronald E. Aubert, PhD

ABSTRACT

Objectives. Although educational achievement is positively related to levels of high-density lipoprotein cholesterol (HDL-C) among White adults, there is an inverse association among Blacks. We assessed whether this interaction could be attributed to differences in the relation of education to correlates of HDL-C.

Methods. Cross-sectional analyses were based on data from 8391 White and 995 Black adults who participated in the Second National Health and Nutrition Examination Survey.

Results. Associations between education and HDL-C levels varied from negative (Black men), to nearly nonexistent (White men and Black women), to positive (White women). Mean HDL-C levels were higher among Blacks than among Whites, but differences varied according to educational achievement. Among adults with less than 9 years of education, mean levels were 6 to 10 mg/dL higher among Blacks, but the racial difference was less than 1 mg/dL among adults with at least 16 years of education. About 20% to 40% of these differences could be accounted for by obesity, alcohol consumption, and other characteristics.

Conclusions. Because of the implications for coronary heart disease risk, consideration should be given to behavioral characteristics associated with the interaction between race and educational achievement. (*Am J Public Health.* 1992;82:999–1006)

Introduction

Rates of coronary heart disease (CHD) in the United States are no higher among Black men than among their White counterparts[1–4] despite the higher prevalence of hypertension among Blacks.[5] These paradoxical findings may be due to differences in levels of high-density lipoprotein cholesterol (HDL-C): mean levels are about 20% (7 to 11 mg/dL) higher among Black men than among White men but are more similar between Black and White women.[4,6–9] Low levels of HDL-C have been shown, in several studies of White populations,[10] to increase the risk for CHD and are associated with more severe coronary artery disease among Black men.[11,12]

Levels of HDL-C are related positively to alcohol consumption and physical activity and inversely to obesity and cigarette smoking, but most information has been obtained from studies of White populations.[13,14] Although HDL-C levels among Blacks are also related to alcohol consumption (positively) and Quetelet index (overweight) (negatively),[4,7,9] associations with socioeconomic status appear to differ between races. Whereas a positive association between HDL-C and various indicators of socioeconomic status is typically observed among Whites,[4,7,15] an inverse association has been reported among Blacks.[4,7,9,16]

The association between socioeconomic status and levels of HDL-C may, in part, be mediated by behavioral characteristics.[15] In agreement with this hypothesis, we have found that the Black/White difference in the association between educational achievement and levels of HDL-C among 31- to 45-year-old male veterans could, in part, be attributed to relative weight, cigarette smoking, and alcohol consumption.[9] The objectives of the current study are to (1) assess these findings in an independent sample, (2) extend these observations to women and to older men, and (3) incorporate information on additional characteristics such as physical activity and body fat distribution. Data are from a representative sample of US adults who participated in the Second National Health and Nutrition Examination Survey (NHANES II).

Materials and Methods

Study Population

Information concerning various risk factors for CHD and the sample of adults with HDL-C determinations in NHANES II, a multistage survey, has been described.[7] Data were collected between 1976 and 1980 by the National Center for Health Statistics through standardized interviews and physical examinations.

Of the 20- to 74-year-olds comprising the original sample of 17 058 persons, 88% (n = 15 080) were interviewed and 68% (n = 11 637) were interviewed and examined. An investigation of possible selection biases[17] indicated that nonresponse at the interview stage did not vary according to the sex or race of the head of household,

David S. Freedman, David F. Williamson, and Ronald E. Aubert are with the Centers for Disease Control's Center for Chronic Disease Prevention and Health Promotion in Atlanta, GA. David S. Strogatz is with the New York State Department of Health and the State University of New York at Albany.

Requests for reprints should be sent to David S. Freedman, PhD, Centers for Disease Control, K-26, 1600 Clifton Road, Atlanta, GA 30333.

This paper was submitted to the Journal January 29, 1991, and accepted with revisions October 18, 1991.

but participation was lower for urban subjects than for those in more rural areas (86% vs 93%). Although persons with self-perceived health problems were more likely to be examined, the assessed health status of participants was comparable to that of subjects in the National Health Interview Survey, which had a 4% nonresponse rate.[17] Race, education, and income did not differ between examined and nonexamined persons in NHANES II.

Of the 11 637 adults examined in NHANES II, HDL-C determinations were performed for 83% (n = 9625), yielding an overall response rate of 56.4% of the original sample. The proportion of examined persons with HDL-C levels varied by race (Whites, 83%; Blacks, 78%) but not by sex, educational achievement, income, age, alcohol consumption, cigarette smoking, or several other characteristics.[7]

Risk-Factor Information

We used educational achievement, determined by the highest attended grade, as an indicator of socioeconomic status. As compared with indices such as income and occupation, education remains fairly constant throughout adulthood, is available for nonemployed and retired persons, and is unaffected by illness. HDL-C levels were related similarly to income and educational achievement in NHANES II.[7]

Race was determined by interviewers' observation or direct questioning. About 4% (n = 339) of White adults indicated they were of Hispanic descent, with Mexicans and Mexican Americans representing the largest subgroup (n = 196). Although most analyses did not distinguish between Hispanic Whites and non-Hispanic Whites, the former were excluded from certain sensitivity analyses.

Quetelet index (kg/m^2) and skinfold thickness measurements were used as measures of overweight and obesity, respectively. The ratio of subscapular to triceps skinfold thickness, an index of central (truncal) fat, was used to describe body fat distribution. Interviewees were asked two questions concerning physical activity: (1) were they getting "much," "moderate," or "little or no" recreational exercise? and (2) aside from recreation, were they usually "very active," "moderately active," or "quite inactive"? These responses were dichotomized as very active/much exercise versus other responses in some analyses.

Self-reported information was used for several prevalent diseases and medications. Diabetes mellitus was defined by positive responses to two questions: "Do you have diabetes or sugar diabetes?" and "Did a doctor tell you this?" Oral contraceptive use was ascertained through the question, "Have you taken birth control pills in the last 6 months and are you taking them now?" Women who reported taking hormones within the last week were considered to be using estrogens. Use of steroids, cholesterol-lowering drugs, and hydantoins was also determined from questions concerning use within the previous week.

Subjects were asked about their usual consumption of beer, wine, and liquor over the previous 3 months, and the number of weekly drinks (summed over the three alcohol categories) was calculated. Cigarette smoking was also assessed, and the reported number of cigarettes smoked daily was used in the analyses. Data on intake of total calories and saturated fats were collected through 24-hour dietary recalls, and subjects were asked if the foods were representative of their usual diet.

Serum Lipid Determinations

About 50% of the examinees were fasting at the time of examination, and in a sample of adults with an oral glucose tolerance test, mean levels of total cholesterol differed by 3 mg/dL or less between fasting and nonfasting persons.[18] Serum was frozen within 1 hour of collection, and specimens were shipped (at approximately 2-week intervals) for lipid analyses to the Lipid Research Clinics Laboratory at George Washington University. Analyses were performed according to standardized protocols,[19] with HDL-C assayed using a heparin and manganese chloride precipitation method.

Statistical Methods

Of the 9625 subjects with HDL-C measurements, persons were excluded if data were missing for one or more of the following characteristics: years of education (n = 85), subscapular (n = 88) or triceps (n = 24) skinfolds, weight or height (n = 51), number of daily cigarettes (n = 14), or alcohol consumption (n = 15). These restrictions resulted in a sample size of 9386 adults (8391 White, 995 Black) for most analyses. Persons were excluded from analyses that used data from the 24-hour dietary recalls if the reported foods did not represent their usual diet (n = 3503) or if their caloric intake was less than 500 kcal or 6000 kcal or more (n = 214).

All analyses incorporate the weights and complex sampling design of NHANES II, with variances calculated using SESUDAAN.[20] After levels of various characteristics between Blacks and Whites were contrasted, associations among various characteristics were examined using linear regression (SUREGR)[21] and stratification. A Race [0 (Whites), 1 (Blacks)] × Education interaction term was incorporated into several analyses as an independent variable. Because education and age were inversely associated, analyses were age adjusted with age treated as a six-level categorical variable: 20–24, 25–34, 35–44, 45–54, 55–64, and 65–74 years.

Several analyses were performed to determine if the observed interaction between race and education could be "explained" by other characteristics. The magnitude of the Race × Education term in regression analyses was compared before and after inclusion of (1) all covariates and (2) each covariate separately. All *P* values and confidence intervals are two-sided.

Results

Mean levels of age and total cholesterol varied only slightly across race-sex groups, but HDL-C levels were higher among Blacks than among Whites (Table 1), with differences of 7 mg/dL among men and 3 mg/dL among women. Whites were more likely to have had some education beyond high school (e.g., 41% of White men vs 26% of Black men) and an income greater than $20 000 per year.

Race-related differences were also observed for other characteristics. Compared with White women, Black women were more obese and had a more central distribution of body fat, as estimated by the ratio of subscapular to triceps skinfolds. Among men, Blacks also had a more central fat distribution than did Whites despite having a *lower* mean Quetelet index. Blacks were also more likely to smoke cigarettes, but Black smokers reported a smaller number of daily cigarettes than did White smokers. Blacks tended to be more physically active and to have lower daily intakes of total calories and saturated fats than did Whites. Differences in the prevalence of diabetes and in the use of oral contraceptives and sex hormones were also observed.

The relation of educational achievement to levels of HDL-C is shown in Figure 1. (Because education and age were inversely associated [$r = -0.29$ to -0.55], mean levels of HDL-C have been age

TABLE 1—Differences in Mean Levels of Various Characteristics between Blacks and Whites in Sample

	Men		Women	
	Whites (n = 3938)	Blacks (n = 443)	Whites (n = 4453)	Blacks (n = 552)
Age (y)	42 ± 20[a]	−2[b] (−4, −1)	42 ± 25	−3 (−4, −1)
Total cholesterol (mg/dL)	211 ± 75	−5 (−12, 2)	215 ± 97	−4 (−11, 3)
HDL cholesterol (mg/dL)	44 ± 26	+7 (6, 9)	53 ± 28	+3 (1, 5)
Education beyond high school (%)	41	−15 (−24, −6)	32	−9 (−16, −1)
Annual income >$20 000 (%)	33	−15 (−25, −6)	27	−13 (−21, −6)
Quetelet index (kg/m²)	25.4 ± 5	−0.4 (−0.6, −0.1)	24.6 ± 8	+2.5 (1.9, 3.1)
Subscapular skinfold (mm)	18 ± 13	0 (−1, 1)	20 ± 21	+6 (5, 7)
Centrality index	1.4 ± 0.8	+0.1 (0.1, 0.2)	0.8 ± 0.4	+0.2 (0.1, 0.2)
Current drinker (%)	76	−1 (−8, 6)	59	−5 (−15, 5)
Drinks per week[c]	5 ± 9	0 (−1, 1)	3 ± 4	0 (0, 0)
Current smoker (%)	40	+14 (7, 20)	34	+1 (−4, 6)
Cigarettes per day[c]	24 ± 15	−8 (−10, −6)	20 ± 12	−5 (−7, −2)
High physical activity[d]				
Recreational (%)	25	+4 (−2, 10)	13	+5 (1, 9)
Nonrecreational (%)	39	+8 (2, 13)	30	+6 (1, 11)
Medication use				
Oral contraceptives (%)	...	...	11	+3 (1, 6)
Hormones (%)	...	...	8	−5 (−9, −2)
Diabetes (%)	2	+1 (0, 3)	3	+2 (0, 3)
Dietary intake[e]				
Total calories (kcal)	2429 ± 1303	−323 (−493, −153)	1544 ± 817	−125 (−213, −38)
Saturated fats (g)	37 ± 26	−7 (−11, −4)	22 ± 15	−3 (−4, −1)
(% of kcal)	13 ± 5	−1 (−2, 0)	13 ± 5	0 (−1, 0)

[a]Values are mean ± standard deviations; the complex sampling design was accounted for in the calculation of the standard deviations.
[b]Values are Black–White difference in mean levels, with 95% confidence interval (CI) shown in parentheses. A positive value indicates that mean levels are higher among Blacks than among Whites.
[c]Restricted to current drinkers (n = 5677) or current smokers (n = 3221).
[d]High (or "much") category versus moderate/little.
[e]Restricted to the 5774 (2567 White men, 274 Black men, 2613 White women, and 320 Black women) subjects who reported consuming their usual diet.

adjusted.) Among White men, mean HDL-C levels varied only slightly with education, but there was a 7 mg/dL decrease over the five education categories among Black men. The relation also differed between Black and White women, with a positive association seen among Whites only. Within each sex, Black/White differences in mean levels of HDL-C were less than 1 mg/dL among persons with at least 16 years of education. Additional analyses treating education as a continuous variable indicated that the Black excess in HDL-C levels decreased by 0.6 mg/dL (women) to 0.7 mg/dL (men) with each additional year of education; the Race × Education interaction term was statistically significant among men ($P = .004$) and women ($P = .02$).

Cross-sectional associations between levels of HDL-C and other characteristics are shown in Table 2; values represent the predicted change in HDL-C for a specified change in each variable. For example, with each additional year of education, levels of HDL-C were estimated to change by +0.8 mg/dL among White women, by +0.2 mg/dL among White men and Black women, and by −0.5 mg/dL among Black men. Levels of HDL-C were inversely associated with obesity and, with the exception of Black men, with the amount of centrally located body fat. As expected, HDL-C levels were positively related to alcohol consumption, but an inverse association with cigarette smoking was seen among Whites only. Physical activity tended to be positively related to levels of HDL-C. Neither total calories nor saturated fats was associated with levels of HDL-C among men, but weak, positive associations were seen in women.

Several of the dichotomous variables were also associated with HDL-C levels, with higher mean levels seen among hydantoin users and women using sex hormones. Although associations with cholesterol-lowering drugs were inconsistent, the mean level of HDL-C was 17 mg/dL lower among the two Black men who reported using these medications than among other Black men. Mean levels of HDL-C were also lower among diabetics (vs nondiabetics) and among the 10 men who reported using steroids (data not shown).

Associations between education and various characteristics were then examined, and several could possibly account for the interaction between race and education. Among men (Table 3), levels of Quetelet index and subscapular skinfold thickness increased with educational achievement among Blacks but not among Whites. In addition, associations with alcohol consumption, cigarette smoking, dietary intake of total calories and saturated fat, and diabetes also differed between Whites and Blacks. Among men with at least 16 years of education, Blacks were more obese, had a higher prevalence of diabetes, and consumed less alcohol than did Whites.

Among women (Table 4), education showed a stronger (inverse) association with overweight and obesity among Whites than among Blacks. For example, Blacks who had completed up to 8 years of education were, on average, 2 kg/m² heavier than their White counterparts, but the difference increased to 3.9 kg/m² among women with at least 16 years of education. Alcohol intake was positively associated with education among White, but not Black, women. Positive associations with saturated fat intake and use of oral contraceptives and sex hormones were stronger among Blacks than among Whites.

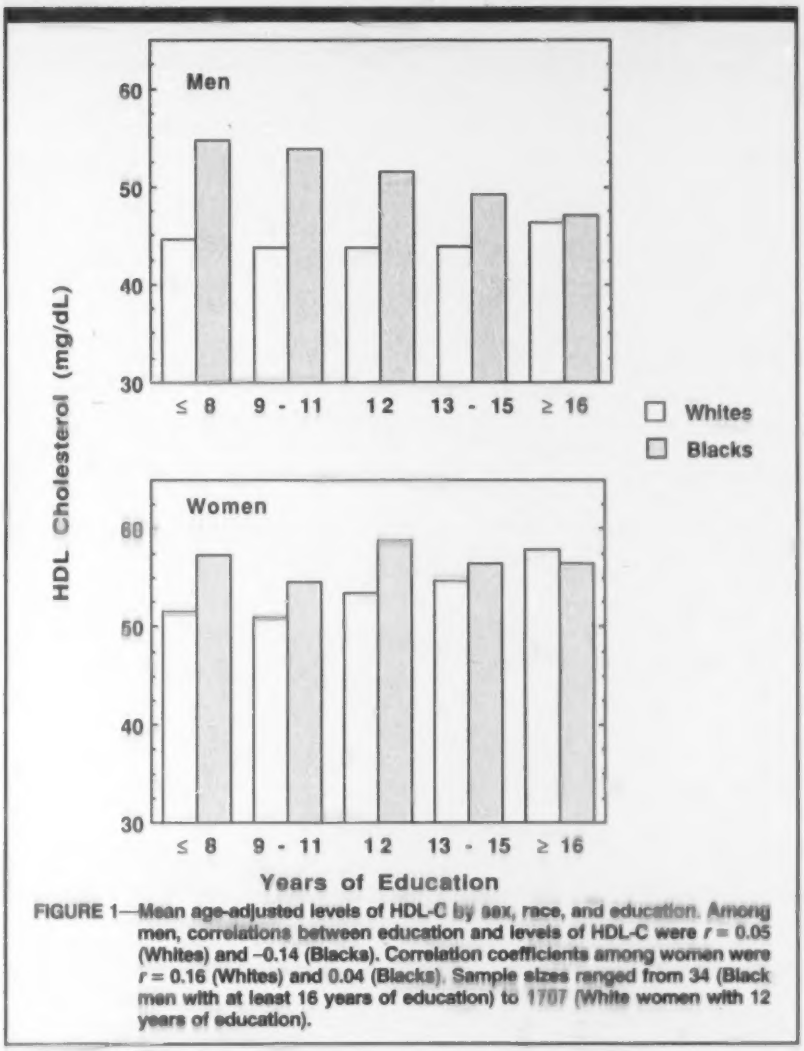


FIGURE 1—Mean age-adjusted levels of HDL-C by sex, race, and education. Among men, correlations between education and levels of HDL-C were $r = 0.05$ (Whites) and –0.14 (Blacks). Correlation coefficients among women were $r = 0.16$ (Whites) and 0.04 (Blacks). Sample sizes ranged from 34 (Black men with at least 16 years of education) to 1707 (White women with 12 years of education).

Analyses were then performed to determine whether the interaction between race and education was attributable to the covariates. Coefficients and predicted differences (Black–White) in HDL-C levels from two regression models are shown in Table 5. Independent variables in model #1 include race, education, race × education, and age; model #2 contains, in addition to these variables, obesity, alcohol consumption, cigarette smoking, and other characteristics. Adjusting for the additional covariates reduced the coefficient of the race × education interaction term by 19% (from –0.72 [model #1] to –0.58 [model #2]) among men and by 41% (from –0.56 to –0.33) among women. Age-stratified analyses indicated that the covariates were important among both younger and older adults (less than 45 years old and 45 to 74 years old, respectively) in accounting for the Race × Education interaction. Most of the variability in levels of HDL-C remained unexplained, however, with the multiple correlation coefficient (R^2) reaching a maximum of 0.18.

Additional analyses, comparable to those shown in Table 5, were performed to examine the effects of various exclusions and individual variables (data not shown). After the 339 Hispanic Whites were excluded, reductions in the magnitude of the Race × Education interaction between the two models were similar to those shown in Table 5. As assessed by analyses in which each characteristic was added separately to model #1, subscapular skinfold thickness and alcohol consumption were the most important covariates in accounting for the observed interaction. Furthermore, covariate adjustment reduced the absolute magnitude of the association between levels of HDL-C and education within each race-sex group. For example, the predicted change in levels of HDL-C with each year of education was reduced from +0.8 to +0.5 mg/dL (White women) and from –0.5 to –0.3 mg/dL (Black men) after controlling for alcohol consumption and subscapular skinfold thickness.

Discussion

Our results indicate that differences between Blacks and Whites in the cross-sectional association between educational achievement and levels of HDL-C can, in part, be attributed to behavioral characteristics. Consistent with the results of other investigations,[4,6,9] mean levels of HDL-C were 3 mg/dL (women) to 7 mg/dL (men) higher among Blacks than among Whites. The largest Black/White differences in HDL-C levels were seen among adults with up to 8 years of education, whereas no difference was evident among persons with at least 16 years of education. With each additional year of education, mean HDL-C levels decreased by 0.5 mg/dL among Black men but not among White men ($\beta = +0.2$ mg/dL per year of educational achievement); among women, the association was positive among Whites ($\beta = +0.8$) but very weak among Blacks ($\beta = +0.2$). About 20% (men) to 40% (women) of the interactions between race and education could be accounted for by other correlates of HDL-C levels, particularly obesity and alcohol consumption.

Although only 56% of the original sample for NHANES II had an HDL-C determination, there were no differences in (1) interview status according to race or sex of the head of household,[17] (2) participation in the examinations by education or income,[17] or (3) obtainment of an HDL-C measurement according to levels of education, Quetelet index, alcohol consumption, or smoking status.[7] It is possible, however, that associations among race, educational achievement, HDL-C levels, and the explanatory variable may be biased by other factors related to participation. The similarity of many of our findings with those from other studies, however, makes this less likely.

With increasing industrialization, the burden of CHD appears to shift to lower social strata,[22] and persons with high socioeconomic status are no longer at the greatest risk.[1,23–26] It has been suggested[27] that these temporal changes occur because persons with high socioeconomic status can more quickly adopt detrimental behaviors that influence CHD risk and are

TABLE 2—Cross-Sectional Associations between Levels of HDL-C and Various Characteristics

	Predicted Change[a] in HDL-C (mg/dL)			
	White Men (n = 3936)	Black Men (n = 443)	White Women (n = 4453)	Black Women (n = 562)
Education (1 y)[b]	0.2[c]	–0.5	0.8	0.2[c]
Quetelet Index (1 kg/m²)	–0.7	–1.3	–0.7	–0.8
Subscapular skinfold (1 mm)	–0.3	–0.5	–0.3	–0.4
Centrality Index (0.5 unit)	–1.3	0.3[c]	–4.7	–4.2
Alcohol (1 drink per week)	0.4	0.4	1.2	0.6
Cigarettes (1 pack per day)	–1.2	2.2[c]	–2.7	0.9[c]
Physical activity[d]				
Recreational (1 unit)	1.2	1.3[c]	1.6	1.9[c]
Nonrecreational (1 unit)	–0.3[c]	2.1	0.6[c]	2.9
Dietary intakes				
Total calories (500 kcal)	0.1[c]	0.9[c]	0.7	0.9[c]
Saturated fat (10 g)	–0.1[c]	–0.3[c]	0.5	1.6[c]
(3% of kcal)	0.0[c]	–1.4	0.1[c]	1.4
Use of				
Hydantoins	11.7 (14)[e]	–2.8[c](3)	6.3[c] (18)	18.5[c] (5)
Hormones	...	...	4.2 (354)	2.1[c](15)
Oral contraceptives	...	...	1.1[c](384)	–1.8 (56)
Cholesterol-lowering drugs	–3.0[c] (20)	–17.0[c](2)	4.3[c] (28)	... (0)

Note. All *P* values are statistically significant at the .05 level except where indicated: [c]*P* > .05.
[a]Change was estimated by linear regression models in which HDL-C was the dependent variable; independent variables include specified characteristic and age.
[b]Values in parentheses represent incremental change in the independent variable.
[c]Predicted change in levels of HDL-C associated with specified increment in the independent variable; associations have been age adjusted.
[d]Physical activity was assessed using a 3-point scale ranging from 1 = little activity to 3 = much activity.
[e]For dichotomous variables, the number of subjects with the characteristic is shown in parentheses.

TABLE 3—Mean Age-Adjusted Levels of Selected Characteristics According to Educational Achievement in Men

		Years of Education					Correlation Coefficient with Education[a]
	Race	≤8	9–11	12	13–15	≥16	
N	Whites	796	569	1243	619	721	...
	Blacks	147	102	107	53	34	...
Quetelet Index (kg/m²)	Whites	25.3	25.9	25.9	25.7	25.3	–0.01
	Blacks	24.5	25.7	25.6	25.6	26.3	0.10
Subscapular skinfold (mm)	Whites	17	18	19	19	17	0.02
	Blacks	16	19	20	19	21	0.17
Centrality Index	Whites	1.5	1.5	1.5	1.5	1.4	–0.06
	Blacks	1.6	1.7	1.6	1.6	1.5	–0.04
Alcohol intake (drinks/week)	Whites	3	4	4	4	5	0.07
	Blacks	4	3	5	4	2	–0.02
Cigarette smoking number/per day)	Whites	11	12	11	8	5	–0.14
	Blacks	8	9	10	8	6	–0.02
Physical activity[b]							
Recreational	Whites	1.8	1.9	2.0	2.0	2.0	0.08
	Blacks	1.8	1.9	1.9	1.9	2.0	0.02
Nonrecreational	Whites	2.3	2.4	2.3	2.2	2.0	–0.15
	Blacks	2.3	2.4	2.4	2.3	2.1	–0.07
Dietary intakes							
Total calories (kcal)	Whites	2207	2339	2388	2415	2181	–0.02
	Blacks	1870	1941	2269	2271	2163	0.12
Saturated fat (% of kcal)	Whites	14	13	13	14	13	–0.03
	Blacks	12	12	13	14	14	0.10
Diabetes (%)	Whites	4	2	3	3	3	...[c]
	Blacks	5	4	5	5	12	...

[a]All correlations have been adjusted for the effects of age.
[b]Physical activity is coded as a 3-level ordinal variable in which 1 = little, 2 = moderate, and 3 = much.
[c]Correlation is not given because characteristic is a dichotomous variable.

TABLE 4—Mean Age-Adjusted Levels of Selected Characteristics According to Educational Achievement in Women

	Race	Years of Education ≤8	9–11	12	13–15	≥16	Correlation Coefficient with Education[a]
N	Whites	774	711	1707	710	551	...
	Blacks	171	126	155	57	43	...
Quetelet index (kg/m²)	Whites	26.5	25.9	25.0	24.3	23.3	–0.16
	Blacks	28.5	28.3	27.3	26.0	27.2	–0.08
Subscapular skinfold (mm)	Whites	23	22	21	20	17	–0.13
	Blacks	27	29	27	25	28	–0.02
Centrality index	Whites	0.9	0.8	0.8	0.8	0.8	–0.08
	Blacks	1.0	1.0	1.0	0.9	1.0	–0.05
Alcohol intake (drinks/week)	Whites	1	1	2	2	3	0.20
	Blacks	1	1	1	1	2	–0.01
Cigarette smoking (number per day)	Whites	6	9	7	5	3	–0.10
	Blacks	4	7	4	5	3	–0.01
Physical activity[b]							
Recreational	Whites	1.5	1.7	1.7	1.8	1.8	0.13
	Blacks	1.6	1.5	1.7	1.9	2.1	0.15
Nonrecreational	Whites	2.1	2.2	2.2	2.2	2.2	0.04
	Blacks	2.1	2.0	2.3	2.4	2.4	0.11
Dietary intakes							
Total calories (kcal)	Whites	1411	1458	1522	1565	1578	0.07
	Blacks	1278	1412	1445	1403	1336	0.07
Saturated fat (% of kcal)	Whites	13	13	12	12	12	–0.06
	Blacks	12	12	12	12	14	0.09
Diabetes (%)	Whites	6	4	4	4	3	...[c]
	Blacks	7	4	8	5	3	...
Use of							
Oral Contraceptives (%)	Whites	7	8	8	7	8	...
	Blacks	9	8	3	21	20	...
Hormones (%)	Whites	6	8	10	10	11	...
	Blacks	<1	2	4	5	11	...

[a]All correlations have been adjusted for the effects of age.
[b]Physical activity is coded as a 3-level ordinal variable in which 1 = little, 2 = moderate, and 3 = much.
[c]Correlation is not given because characteristic is a dichotomous variable.

TABLE 5—Regression Analyses Examining the Influence of Covariates on the Interaction between Race and Educational Achievement in the Prediction of HDL-C Levels[a]

Sex	Age Group (y)	Model[c]	R^2	Regression Coefficients[b] Education	Race x Education	Predicted Black Excess in Levels of HDL-C (mg/dL) According to Years of Education 0	12	16
Men	Overall (n = 4374)	1	0.04	0.2	–0.72	16.2	7.8	4.7
		2	0.14	0.1	–0.58	14.3	7.3	5.0
	20–44 (n = 1982)	1	0.05	0.4	–0.66	15.5	7.7	5.0
		2	0.13	0.2	–0.37	11.1	6.7	5.3
	45–74 (n = 2392)	1	0.04	0.0	–0.70	15.7	7.3	4.5
		2	0.18	0.0	–0.55	14.6	8.0	5.8
Women	Overall (n = 4999)	1	0.03	0.8	–0.56	10.6	3.9	1.6
		2	0.15	0.3	–0.33	9.4	5.4	4.1
	20–44 (n = 2260)	1	0.03	1.0	–0.84	12.0	4.3	1.8
		2	0.14	0.4	–0.21	7.3	4.8	3.9
	45–74 (n = 2739)	1	0.02	0.7	–0.59	10.1	2.9	0.5
		2	0.17	0.3	–0.27	9.5	6.2	5.1

Note. R^2 = multiple correlation coefficient.
[a]All analyses are restricted to subjects with complete information for all characteristics; eight men and six women have been excluded.
[b]Race is coded as 0 (Whites) or 1 (Blacks) in the interaction term; education is coded as an ordinal variable with 18 levels. The regression coefficient for race is indicated by the Black excess in HDL-C levels for adults with 0 years of education.
[c]The dependent variable in all analyses is HDL-C. Model 1 contains race, education, race × education, and age as predictor variables. Model 2 contains, in addition to the variables included in model 1, Quetelet index; subscapular skinfold thickness; alcohol consumption; cigarette smoking; work-related physical activity; nonwork-related physical activity; diabetes; and use of hydantoins, cholesterol-lowering drugs, hormones (women only), oral contraceptives (women only), and steroids (men only).

then able to modify these behaviors more rapidly. Consistent with this hypothesis, the relation of socioeconomic status to levels of total cholesterol, cigarette smoking, and consumption of various foods has changed between the 1950s and 1980s.[23,28] Inverse associations between indicators of socioeconomic status and several CHD risk factors have been found,[15,23,28–30] and a positive association between educational achievement and levels of HDL-C has been reported in studies of White adults.[15,30]

In contrast, socioeconomic status was recently found to be *positively* related to CHD incidence among urban men in Puerto Rico, possibly due to the relatively recent economic development of this area.[31] Even within industrialized societies, however, the relation of socioeconomic status to CHD may vary across ethnic groups.[32] A study from one center in the Multiple Risk Factor Intervention Trial, for example, found that socioeconomic status was inversely related to a composite risk-factor score for CHD among White, Asian, and Hispanic men, but not among Black men.[33] Furthermore, a recent survey of adults from New York City found an inverse association between educational achievement and several CHD risk factors (smoking, sedentary behavior, obesity, and atherogenic diet) that was stronger among Whites than among Blacks.[34]

Our findings concerning the inverse association between educational achievement and levels of HDL-C among Black men agree with the results of other investigations. Mean HDL-C levels among Black adults from Framingham, Mass (50% of whom had completed college), were 6 to 7 mg/dL lower than those among Whites.[35] Black/White differences in the association between socioeconomic status indicators and levels of HDL-C have also been found in the Multiple Risk Factor Intervention Trial[4] and among US Army veterans.[9] With the exception of education, we found that associations between levels of HDL-C and most characteristics did not differ between Blacks and Whites. Although cigarette smoking was inversely associated with levels of HDL-C among Whites only, other studies of smoking and HDL-C levels in Blacks have yielded conflicting results, with the observed associations ranging from inverse[4] to weakly positive.[35]

Few studies have examined characteristics that might mediate the relation of educational achievement to levels of HDL-C. Heiss et al.,[15] however, in an analysis of White adults, found that most of the positive association between education and HDL-C levels among women could be attributed to obesity, alcohol consumption, smoking, and use of sex hormones. (Associations between education and these characteristics were similar to those seen in the current study.) Furthermore, about 40% of the interaction between race and educational achievement for levels of HDL-C could be attributed to Quetelet index, alcohol consumption, and cigarette smoking among 31- to 45-year-old men.[9] Our results, based on a national probability survey, extend these findings to women and older men. The persistence of a large proportion of the interaction even after covariate adjustment could result from (1) unmeasured characteristics that are associated with levels of HDL-C and educational achievement, or (2) misclassification of HDL-C levels and other characteristics. Because of the large day-to-day variability in dietary intake, our assessed intakes of total calories and saturated fat may be particularly inaccurate.

When one considers the differences in job opportunities, income, and physical environment that exist between Blacks and Whites at any specified level of education,[36] racial differences in the association between socioeconomic status and measures of health and health-related behaviors[37–40] are not surprising. A full explanation of Black/White differences in the association between educational achievement and levels of HDL-C remains to be resolved. Our findings, however, suggest that preventing obesity may be a useful goal for maintaining the relatively high levels of HDL-C seen among Blacks in the United States. □

Acknowledgments

The authors acknowledge the support of Tim Byers at the Centers for Disease Control. We also thank Robinson Fulwood, of the National Heart, Lung, and Blood Institute, and Margaret Carroll, of the National Center for Health Statistics, for their advice and encouragement in analyzing these data.

References

1. Cassel J, Heyden S, Bartel AG, et al. Incidence of coronary heart disease by ethnic group, social class, and sex. *Arch Intern Med.* 1971;128:901–906.
2. Gillum RF. Coronary heart disease in Black populations: I. mortality and morbidity. *Am Heart J.* 1982;104:839–851.
3. Keil JE, Loadholt CB, Weinrich MC, et al. Incidence of coronary heart disease in Blacks in Charleston, South Carolina. *Am Heart J.* 1984;108(pt 2):779–786.
4. Watkins LO, Neaton JD, Kuller LH. Racial differences in high-density lipoprotein cholesterol and coronary heart disease incidence in the usual-care group of the Multiple Risk Factor Intervention Trial. *Am J Cardiol.* 1986;57:538–545.
5. Gillum RF. Coronary heart disease in Black populations: II. risk factors. *Am Heart J.* 1982;104:852–864.
6. Tyroler HA, Glueck CJ, Christensen B, Kwiterovich PO. Plasma high-density lipoprotein cholesterol comparisons in Black and White populations: the Lipid Research Clinics Program Prevalence Study. *Circulation.* 1980;62(suppl IV):99–107.
7. Linn S, Fulwood R, Rifkind B, et al. High-density lipoprotein cholesterol levels among US adults by selected demographic and socioeconomic variables: the Second National Health and Nutrition Examination Survey 1976–1980. *Am J Epidemiol.* 1989;129:281–294.
8. Sprafka JM, Folsom AR, Burke GL, et al. Prevalence of cardiovascular disease risk factors in Blacks and Whites: the Minnesota Heart Survey. *Am J Public Health.* 1988;78:1546–1549.
9. Freedman DS, Strogatz DS, Eaker E, Joesoeph MR, DeStefano F. Differences between Black and White men in correlates of HDL cholesterol. *Am J Epidemiol.* 1990; 132:656–669.
10. Gordon DJ, Probstfield JL, Garrison RJ, et al. High-density lipoprotein cholesterol and cardiovascular disease: four prospective American studies. *Circulation.* 1989;79:8–15.
11. Cooper R, Sempos C, Ghali J, et al. High-density lipoprotein cholesterol and angiographic coronary artery disease in Black patients. *Am Heart J.* 1985;110:1006–1011.
12. Freedman DS, Gruchow HW, Manley JC, et al. Black/White differences in risk factors for arteriographically-documented coronary artery disease. *Am J Cardiol.* 1988;62:214–219.
13. Heiss G, Johnson NJ, Reiland S, et al. The epidemiology of plasma high-density lipoprotein cholesterol levels: the Lipid Research Clinics Program Prevalence Study. *Circulation.* 1980;62(suppl IV):116–136.
14. Wilson PWF, Garrison RJ, Abbott RD, Castelli WP. Factors associated with lipoprotein cholesterol levels: the Framingham Study. *Arteriosclerosis.* 1983;3:273–281.
15. Heiss G, Haskell W, Mowery R, et al. Plasma high-density lipoprotein cholesterol and socioeconomic status: the Lipid Research Clinics Program Prevalence Study. *Circulation.* 1980;62(suppl IV):108–115.
16. Ford E, Cooper R, Simmons B, Katz S, Patel R. Sex differences in high-density lipoprotein cholesterol in urban Blacks. *Am J Epidemiol.* 1988;127:753–762.
17. Forthofer RN. Investigation of nonresponse bias in NHANES II. *Am J Epidemiol.* 1983;117:507–515.
18. National Center for Health Statistics, National Heart, Lung, and Blood Institute Collaborative Lipid Group. Trends in serum cholesterol levels among US adults aged 20 to 74 years. *JAMA.* 1987;257:937–942.
19. *Manual of Laboratory Operation: Lipid Research Clinics Program.* Vol 1. Bethesda, Md: National Institutes of Health;

1974. US Dept of Health, Education, and Welfare publication NIH 75-628.
20. Shah BV. *SESUDAAN: Standard Errors Program for Computing of Standardized Rates from Sample Data.* Research Triangle Park, NC: Research Triangle Institute; 1981.
21. Shah BV. *SURREGR: Standard Errors of Regression Coefficients from Sample Survey Data.* Research Triangle Park, NC: Research Triangle Institute; 1982.
22. Jenkins CD. Recent evidence supporting psychologic and social risk factors for coronary disease. *N Engl J Med.* 1976;294: 987–994.
23. Marmot MG, Adelstein AM, Robinson N, et al. Changing social-class distribution of heart disease. *Br Med J.* 1978;2:1109–1112.
24. Rose G, Marmot MG. Social class and coronary heart disease. *Br Heart J.* 1981;45: 13–19.
25. Rosenman RH, Brand RJ, Jenkins CD, et al. Coronary heart disease in the Western Collaborative Group Study: final follow-up experience of 8½ years. *JAMA.* 1975;233: 872–877.
26. Liu K, Cedres LB, Stamler J, et al. Relationship of education to major risk factors and death from coronary heart disease, cardiovascular disease and all causes: findings of the three Chicago epidemiologic studies. *Circulation.* 1982;66:1308–1314.
27. Morgenstern H. The changing association between social status and coronary heart disease in a rural population. *Soc Sci Med.* 1980;14A:191–201.
28. Stamler J. Primary prevention of coronary heart disease: the last 20 years. *Am J Cardiol.* 1981;47:722–735.
29. Jacobsen BK, Thelle DS. Risk factors for coronary heart disease and level of education: the Tromsø Heart Study. *Am J Epidemiol.* 1988;127:923–932.
30. Helmert U, Herman B, Joeckel KH, Greiser E, Madans J. Social class and risk factors for coronary heart disease in the Federal Republic of Germany: results of the baseline survey of the German Cardiovascular Prevention Study (GCP). *J Epidemiol Community Health.* 1989;43: 37–42.
31. Sorlie PD, Garcia-Palmieri MR. Educational status and coronary heart disease in Puerto Rico: the Puerto Rico Heart Health Program. *Int J Epidemiol.* 1990;19: 59–65.
32. Marmot MG, Kogevinas M, Elston MA. Social/economic status and disease. *Ann Rev Public Health.* 1987;8:111–135.
33. Kraus JF, Borhani NO, Franti CE. Socioeconomic status, ethnicity, and risk of coronary heart disease. *Am J Epidemiol.* 1980; 111:407–414.
34. Shea F, Stein AD, Basch CE, et al. Independent associations of educational attainment and ethnicity with behavioral risk factors for cardiovascular disease. *Am J Epidemiol.* 1991;134:567–582.
35. Wilson PWF, Savage DD, Castelli WP, et al. HDL-cholesterol in a sample of Black adults: the Framingham Minority Study. *Metabolism.* 1983;32:328–332.
36. US Bureau of the Census. *Statistical Abstract of the United States: 1982–1983.* Washington, DC: US Govt Printing Office; 1984.
37. Kessler RC, Neighbors HW. A new perspective on the relationships among race, social class, and psychological distress. *J Health Soc Behav.* 1986;27:107–115.
38. Strogatz DS. Use of medical care for chest pain: differences between Blacks and Whites. *Am J Public Health.* 1990;80:290–294.
39. Cooper R. A note on the biologic concept of race and its application in epidemiologic research. *Am Heart J.* 1984;108:715–723.
40. Caspar ML, Wing S, Strogatz DS. Variation in the magnitude of Black-White differences in stroke mortality by community occupational structure. *J Epidemiol Community Health.* In press.

Call for Abstracts for Primary Care Nursing Symposium

The 18th National Primary Care Nurse Practitioner Symposium is accepting abstracts of research, innovative clinical practice/models and roles, and other professional papers relating to primary care nursing. Sponsored by the University of Colorado Health Sciences Center School of Nursing, this symposium will be held July 8–11, 1993, at Keystone Resort, Keystone, Colorado.

Preference will be given to abstracts with strong clinical applicability, innovative role-development models, and a future-oriented perspective. Abstracts are being accepted for oral or poster presentation. Deadline for submission is *September 14, 1992.* Contact Ellen Lemberg, RN, MS, Nurse Practitioner Symposium Office, University of Colorado Health Sciences Center School of Nursing, 4200 East 9th Ave, Box C287, Denver, CO 80262; tel. (303) 270-7436.

Young Women's Antisocial Behavior and the Later Emotional and Behavioral Health of Their Children

Sandra L. Martin, PhD, and Margaret R. Burchinal, PhD

ABSTRACT

Objectives. Antisocial and criminal behavior are increasing among US youth, especially among females. Thus, questions arise concerning whether youthful female deviant behavior eventually will have negative behavioral and emotional consequences for the later children of these women.

Methods. Associations between the severity of early female antisocial behavior (including both drug-related and non–drug-related offenses) and the later behavioral and emotional health of the children of these women were examined among 1425 mother–child pairs of the National Longitudinal Survey of Youth.

Results. Multiple linear regression procedures indicated a significant positive relationship between the severity of the mothers' early non–drug-related offenses and the later severity of the children's scores on the Antisocial, Hyperactive, Anxious/Depressed, Headstrong, Peer Conflict/Social Withdrawal, Immature Dependency, and Total Problem subscales of the Behavior Problem Index.

Conclusions. This study demonstrated an association between the antisocial behavior of female youth and the later behavioral and emotional problems of the children of these women. Future research needs to determine the mechanisms underlying the intergenerational transmission of these types of problems so that effective preventive and therapeutic public health practices may be designed and implemented. (*Am J Public Health.* 1992;82:1007–1010)

Introduction

Antisocial and criminal behavior are on the upswing among US youth, especially among females.[1] This trend raises questions concerning whether youthful female deviant behavior will eventually have negative consequences for the later children of these women. Although ties have been established between female antisocial behavior and deviancy in the children of these women, research has tended to focus on the most antisocial of women, in particular, women who have been arrested or incarcerated. For example, studies have found that juvenile delinquents are more likely than nondelinquents to have mothers who have been involved with the law,[2–6] and a high prevalence of school behavior problems have been found in children of incarcerated women.[7] Furthermore, past research concerning the children of antisocial women has usually examined a restricted range of childhood problems, typically juvenile delinquency, ignoring other forms of childhood emotional and behavioral problems. In addition, these studies often have used small clinic samples of subjects rather than large samples of the general population.

In this study, we examine the association between the severity of early female antisocial behavior (including both drug-related and non–drug-related offenses) and the later behavioral and emotional health of the children of these women using a secondary data analysis of information from a nationwide prospective study of American youth, specifically, the National Longitudinal Survey of Youth (NLSY).

Materials and Methods

The National Longitudinal Survey of Youth

The NLSY began in 1979 as a prospective study of factors that influence the labor force participation of American young people.[8] The original cohort included both men and women, aged 14 to 24 years, with an overrepresentation of African Americans, Hispanics, and economically disadvantaged non-Hispanic Whites. Since the inception of the project, 1-hour personal interviews of cohort members have been conducted annually, with 95% of the sample being reinterviewed each year.

Interview content varies from year to year, with information being collected on a multitude of factors, including socioeconomic, sociodemographic, familial, health, and behavioral characteristics. During the 1980 interview, the Self Reported Delinquency Index[9] was used to assess the frequency of cohort participation in antisocial activities (e.g., vandalism, selling drugs). During the 1986 interview, female participants who were mothers of children at least 4 years of age were administered the Behavior Problem Index[10,11] to assess their children's emotional and behavioral health. Thus, linking the 1980 and 1986 NLSY interviews allows examination of the relationship between the antisocial behavior of female youth and the later emotional and behavioral health of the children of these women.

Sandra L. Martin is with the Department of Maternal and Child Health, and Margaret R. Burchinal is with the Department of Psychology and the Frank Porter Graham Child Development Center, University of North Carolina at Chapel Hill.

Requests for reprints should be sent to Sandra L. Martin, PhD, Department of Maternal and Child Health, The School of Public Health, University of North Carolina, Chapel Hill, NC 27599-7400.

This paper was submitted to the Journal May 8, 1991, and accepted with revisions November 12, 1991.

TABLE 1—Number and Percentage of Women Who Committed Each Type of Offense and the Mean Annual Frequency of Each Offense

Offense	Subjects Committing Offense		Offense Frequency	
	n	%	Mean	SD
Drug				
Using hard drugs	232	16	10.53	14.67
Selling marijuana	75	5	10.72	15.78
Selling hard drugs	19	1	12.79	17.92
Nondrug				
Assault	419	29	4.91	9.99
Shoplifting	320	22	3.98	8.28
Physically fighting at school/work	272	19	2.71	4.66
Fraud	264	19	5.16	9.59
Petty theft	132	9	2.80	5.52
Vandalism	110	8	2.35	3.74
Aggravated assault	108	8	2.81	5.95
Fencing stolen goods	83	6	3.34	7.68
Auto theft	72	5	3.96	8.07
Robbery	35	2	4.11	9.96
Grand theft	21	1	1.52	1.54
Break and enter	19	1	4.47	11.16
Run gambling operation	19	1	11.53	18.80

Sample Selection

Of the 1612 NLSY women who participated in the assessment of their children's behavioral and emotional health, 1425 (88%) had complete information available for all selected analysis variables for both the 1980 and 1986 interviews; therefore, these 1425 women and their first-born children were selected to be included in this study. Comparison of the 1425 women with complete information with the 187 women with incomplete information indicated no statistically significant differences ($P < .05$) in terms of the women's ethnicity, age, educational attainment, employment histories, marital status, number, age and gender of children, and children's scores on the emotional/behavioral assessments.

Assessment of the Women

The severity of the antisocial behavior of the women in 1980 was assessed using the Self Reported Delinquency Index.[9] This instrument is a listing of 16 specific antisocial behaviors (e.g., stealing items worth more than $50, selling hard drugs) to which the subject responds by noting the frequency of his or her participation in each of the behaviors during the past year (i.e., "never," "once," "twice," "three to five times," "six to ten times," "eleven to fifty times," and "more than 50 times"). Item scores are calculated by assigning the midpoint value of the first five response categories and assigning a score of 50 to those who responded in the "50 or more" category. Then, two summary indices, severity of drug offenses and severity of nondrug offenses, are created by summing the subject's scores to pertinent behavioral items and dividing by the number of items composing each index (see Table 1 for the offenses included in each summary index). For analysis purposes, a square root transformation was used on the severity indices to help suppress the effects of the few extreme outlying values of these predictor variables.

Care was taken in administering the Self Reported Delinquency Index to promote an environment in which the respondents would feel comfortable answering potentially sensitive questions (e.g., respondents were assured of confidentiality; respondents recorded their answers on separate response sheets rather than responding aloud to questions; respondents placed their completed response sheets within envelopes, which they then sealed). Despite the delicate nature of these questions, more than 98% of all subjects completed the assessment. Although it should be recognized that any self-report measure of antisocial behavior is likely to represent an underreport of the true prevalence of the behavior, studies of the Self Reported Delinquency Index, which compared the severity indices with information concerning police contact and school behavior problems, report good reliability and validity.[9]

The 1980 and 1986 interviews also collected information concerning the women's ethnicity, age, years of education, marital status, and poverty status. The poverty status of the women was assessed using modified federal guidelines based on income, family size, and residence.[12,13]

Assessment of the Children's Emotional and Behavioral Health

The emotional and behavioral health of the children was assessed in 1986 using the Behavior Problem Index (BPI).[10,11] This instrument is a listing of 28 specific childhood problem behaviors or symptoms, many items being drawn from previous scales.[14–17] The child's mother responds to each of the BPI items indicating whether the symptom in question has been typically exhibited by her child within the past 3 months. Positive responses to specific items are summed to create a Total Problem score and the following six empirically derived subscale scores: (1) Antisocial, (2) Hyperactive, (3) Anxious/Depressed, (4) Headstrong, (5) Peer Conflict/Social Withdrawal, and (6) Immature Dependency. The Total Problem score and the subscale scores have been standardized by age and gender based on general population samples (mean = 100, SD = 15). Reliability estimates of the Total Problem score range from .86 to .89, while those for the various subscale scores range from .54 to .73.[10]

Analysis

Linear regression analyses[18] were conducted to examine the potential associations between the mothers' severity of drug and nondrug offenses and their children's levels of emotional and behavioral problems. First, the children's BPI Total Problem scores were analyzed. Then, each of the five BPI subscale scores was analyzed. Three sets of predictor variables were used in the models. The first set consisted of the variables of interest, namely, the mother's severity of drug and nondrug offenses indices. The second set of predictors included sampling variables that had been used to select the original NLSY cohort, specifically, ethnicity (African American, non-Hispanic White, Hispanic), poverty (yes, no), and an ethnicity by poverty interaction term. The interaction term was necessary because the NLSY oversampled Hispanics, African Americans, and impoverished non-Hispanic Whites. The third set of predictor variables was selected to control for potential confounders. These variables included the mother's age at the time she

was asked about her antisocial behaviors, the marital status of the mother in 1986, the age of the child in 1986, and the gender of the child. F tests (with a type I error rate of .05) were used to evaluate all statistical tests. Effect sizes for the variables of interest were measured using estimated regression coefficients, and 95% confidence intervals were calculated.

Results

Description of the Women in 1980

The study women included 688 (48%) non-Hispanic Whites, 481 (34%) Hispanics, and 256 (18%) African Americans. Of these women, 496 (35%) were living in poverty in 1980. In 1980, the women ranged in age from 15 to 23 years, with the mean age being 19.58 years (SD = 2.03). Their ages at the birth of the target children ranged from 13 to 24 years (mean = 18.13, SD = 2.22). A total of 713 (50%) of the women were never married, 594 (42%) were married, and 118 (8%) were divorced or separated during 1980. The women's educational levels in 1980 ranged from 0 to 16 years, with the mean being 10.88 years (SD = 1.87).

The results of the Self Reported Delinquency Index showed that 257 (18%) of the study women committed at least one drug offense during 1980 (232 [16%] women used hard drugs, 75 [5%] sold marijuana, and 19 [1%] sold hard drugs). Over all 1425 women, the severity of drug offense index ranged from 0 to 45, with a mean of 0.82 (SD = 3.65).

A total of 814 (57%) study women committed at least one nondrug offense, the most common being assault (419 women; 29%), shoplifting (320 women; 22%), physically fighting at school or work (272 women; 19%), and fraud (264 women; 19%). Over all 1425 women, the severity of nondrug offense index ranged from 0 to 12.92, with a mean of 0.40 (SD = 1.08).

Description of the Women and Children in 1986

The poverty, marital, and educational status of the women remained fairly stable from 1980 through 1986. The poverty status of 71% of the study women remained the same from 1980 through 1986, with 37% of the study women living in poverty in 1986. The marital status of 61% of the women remained constant over the two time points, with 352 (25%) women never being married, 774 (54%) being married, and 299 (21%) being separated or divorced during the 1986 interviews. There was a significant correlation between the women's years of education in 1980 and 1986 (r = .88, P < .0001), with the mean education level of the women in 1986 being 11.30 years (SD = 2.01).

No information concerning drug offenses or nondrug offenses was collected during the 1986 assessment; thus, it cannot be determined whether the women's antisocial and drug-related behaviors remained consistent over the 6-year study period.

In 1986, the target children ranged in age from 4 years to 12 years (mean = 7.45, SD = 2.08), with girls and boys being fairly equally represented in the sample (735 boys [52%] and 690 girls [48%]). The children's mean BPI Total Problem score (mean = 108.8, SD = 13.7) and the means of all six BPI subscale scores (i.e., Antisocial [mean = 108.1, SD = 13.7], Hyperactive [mean = 109.2, SD = 13.8], Anxious/Depressed [mean = 106.8, SD = 13.1], Headstrong [mean = 103.6, SD = 13.1], Peer Conflict/Social Withdrawal [mean = 105.8, SD = 12.9], and Immature Dependency [mean = 106.7, SD = 13.2]) were above the general population mean of 100, perhaps reflecting the large proportion of high-risk families in this sample (e.g., financially deprived families).

When the children's BPI Total Problem scores were analyzed, a statistically significant effect was observed for the mother's severity of nondrug offenses, F(1, 1413) = 30.81, P <.0001, with the children's problem behavior increasing with increasing frequencies of the mothers' nondrug offenses; however, the mother's severity of drug offenses was not found to be significantly related to the children's BPI Total Problem scores. When the six BPI subscale scores were modeled, similar results were found in that severity of nondrug offenses was significantly associated with each of the children's BPI subscale scores: Antisocial, F(1, 1413) = 22.07, P <.0001; Hyperactive, F(1, 1413) = 19.83, P <.0001; Anxious/Depressed, F(1, 1413) = 18.13, P <.0001; Headstrong, F(1, 1413) = 17.35, P <.0001; Peer Conflict/Social Withdrawal, F(1, 1413) = 9.77, P = .0018; and Immature Dependency, F(1, 1413) = 7.20, P = .0074. However, the severity of drug offenses was not significantly related to any of the six BPI subscale scores.

TABLE 2—Estimated Regression Coefficients and Associated 95% Confidence Intervals

Behavior Problem Index Score	Regression Coefficients (95% CI)	
	Drug Offenses	Nondrug Offenses
Total Problem	−0.49 (−1.39, 0.41)	4.27 (2.76, 5.78)
Antisocial	−0.02 (−0.87, 0.91)	3.58 (2.09, 5.06)
Hyperactive	−0.37 (−1.27, 0.53)	3.44 (1.93, 4.95)
Anxious/Depressed	−0.38 (−1.24, 0.48)	3.15 (1.70, 4.60)
Headstrong	−0.52 (−1.37, 0.33)	3.03 (1.61, 4.45)
Peer Conflict/Social Withdrawal	−0.41 (−1.27, 0.45)	2.29 (0.85, 3.73)
Immature Dependency	−0.18 (−1.04, 0.72)	2.01 (0.54, 2.94)

Table 2 presents the estimated regression coefficients and 95% confidence intervals associated with the two variables of interest, severity of drug offenses and severity of nondrug offenses, for each of the models of the children's BPI Total Problem and subscale scores. There is no significant association between the mother's severity of drug offenses and any of the children's BPI scores; however, each of the children's BPI scores tends to increase significantly with increasing levels of the mother's severity of nondrug offenses. It is interesting that the largest effect size relating the mother's severity of nondrug offenses and the children's BPI subscale scores occurs in the model of the children's Antisocial scores.

Discussion

In line with previous research that has demonstrated a link between female criminality and deviancy in the children of these women, this study found an association between the antisocial behavior of female youth and the later behavioral and emotional problems of the children of these women. This study extends past research by examining a range of severity of women's antisocial behaviors and a range of children's behavioral and emotional problems. Specifically, the present study found a positive relationship between the severity of antisocial behavior of female youth and the severity of behavioral and emotional problems of the children of these women 6 years later. This relation-

ship was found for non–drug-related offenses but not for drug-related offenses. However, some explanations and cautionary comments are in order.

First, one must interpret the results of this research with caution given the limitations of both the study design and the correlational nature of the study analysis. The observed association between the severity of women's antisocial behavior and the severity of their children's emotional and behavioral problems is likely due to life-style differences, which tend to be established by the teenage years. These lifestyle differences may be due to unobserved familial and social factors, such as family dysfunction, living in a high-crime neighborhood, and so forth. Thus, it is impossible from this study to specify the potential causal chain that connects the life experiences of adolescent females through to the later emotional and behavioral health of the children of these young women. Future studies of this topic need to concentrate on understanding the specific mechanisms underlying the intergenerational transmission of these types of problems.

Second, one must consider the fact that the measure of the women's drug and nondrug offenses and the measure of the children's emotional health occurred at two separate time points, with approximately 6 years between the two assessments. Therefore, we are unsure whether the early antisocial behavior and drug use of the women remained stable over the 6-year period. Furthermore, many unmeasured events might have occurred during that 6-year period that might have affected the children's emotional and behavioral development. In light of these considerations, it is interesting that we still see an association between the early behavior of the women and the later emotional and behavioral health of their children. However, this finding is not totally unexpected since research has suggested some degree of continuity in values and behaviors from adolescence through to early adulthood.

Third, the measures of female antisocial behavior and children's emotional and behavioral health were based on reports of the study women. As such, they are subject to the biases associated with self-report data.

Fourth, it is noteworthy that we did not find an association between the women's severity of early drug-related offenses and their children's later behavioral and emotional health, while we did find such a relationship when the women's non–drug-related offenses were examined. Further study of the data sheds some light on this finding. Of the 1425 study women, 568 reported committing neither drug nor nondrug offenses, 214 reported committing both drug and nondrug offenses, and 600 reported committing exclusively non–drug offenses; however, only 43 reported committing exclusively drug offenses. In other words, most of the drug offenders also committed nondrug offenses, while only about a third of the women who reported nondrug offenses also reported committing drug offenses. Therefore, the ability of this study to detect a potential effect of drug offenses, independent of that of nondrug offenses, may have been somewhat limited by the relatively small size of the drug offenses alone group.

In spite of the aforementioned study limitations, the findings do suggest some interesting paths for future investigation concerning the potential intergenerational transmission of behavioral and emotional problems. Gaining greater insight concerning the mechanisms underlying the relationship between parents' and children's emotional and behavioral health could, we hope, lead to the design and implementation of effective preventive and therapeutic public health policies aimed at enhancing the health of both parents and children. □

Acknowledgments

This research was supported in part by the Frank Porter Graham Child Development Center and the Carolina Consortium on Human Development.

The authors would like to thank Drs. Lawrence Kupper and Elizabeth Robertson for their thoughtful advice and suggestions during the preparation of this paper and Ms. Ronda Music for her technical assistance.

References

1. Adler F. The United States: international concern in light of the American experience. In Adler F, ed. *The Incidence of Female Criminality in the Contemporary World.* New York, NY: New York University Press; 1981:1–13.
2. Glueck S, Glueck ET. *Five Hundred Delinquent Women.* New York, NY: Knopf; 1934:64–73.
3. West DJ, Farrington DP. *Who Becomes Delinquent?* London, England: Heinemann; 1973:26–46.
4. Glueck S, Glueck ET. *Unravelling Juvenile Delinquency.* New York, NY: The Commonwealth Fund; 1950:108–116.
5. Robins LN, West PA, Herjanic BL. Arrests and delinquency in two generations: a study of Black urban families and their children. *J Child Psychol Psychiatry.* 1975;16: 125–140.
6. Robins LN. *Deviant Children Grown Up.* Baltimore, Md: Williams and Wilkins; 1966:159–180.
7. Stanton AM. *When Mothers Go to Jail.* Lexington, Mass: Lexington Books; 1980: 89–95.
8. Olsen RJ. *The NLS Research Project. The National Longitudinal Surveys of Labor Market Experience.* Columbus, Ohio: Center for Human Resource Research, The Ohio State University, 1988:1–5.
9. Crowley JE. Crime and delinquency: Descriptions and distributions. In Borus ME (ed): Pathways to the Future—A Longitudinal Study of Young Americans—Preliminary Report on the 1980 Survey. Columbus, Ohio: Center for Human Resource Research, The Ohio State University; 1981:73–92.
10. Baker PC, Mott FL. *NLSY Child Handbook 1989, A Guide and Resource Document for the National Longitudinal Survey of Youth 1986 Child Data.* Columbus, Ohio: Center for Human Resource Research, The Ohio State University; 1989: 65–69.
11. Peterson JL, Zill N. Marital disruption, parent-child relationships, and behavioral problems in children. *J Marriage Fam.* 1986;48:295–308.
12. US Bureau of the Census. *Money, Income, and Poverty Status of Families and Persons in the US: 1979.* Washington, DC: US Government Printing Office; 1980:3–6. Current Population Report Series/Consumer Income. Series P-60, No. 125.
13. US Bureau of the Census. *Money, Income, and Poverty Status of Families and Persons in the US: 1985.* Washington, DC: US Government Printing Office; 1986:1–4. Current Population Report Series/Consumer Income. Series P-60, No. 154.
14. Achenbach TM, Edelbrock C. Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. *Monogr Soc Res Child Dev.* 1981;46(serial no. 188, 1):1–78.
15. Graham PJ, Rutter ML. The reliability and validity of the psychiatric assessment of the child, II. Interview with the parent. *Br J Psychiatry.* 1968;114:581–592.
16. Rutter M, Tizard J, Whitmore K. *Education, Health and Behavior.* London, England: Longman; 1970.
17. Kellam SK, Branch JD, Agrawal KC, Ensminger ME. *Measuring Mental Health in First-Graders. Mental Health and Going to School: The Woodlawn Program of Assessment, Early Intervention, and Evaluation.* Chicago, Ill: University of Chicago Press; 1975:25–132.
18. Kleinbaum DG, Kupper LL, Muller KE. *Applied Regression Analysis and Other Multivariable Methods.* 2nd ed. Boston, Mass: Kent Publishing Co; 1988:124–143.

Public Health Briefs

Rotating Shift Work, Sleep, and Accidents Related to Sleepiness in Hospital Nurses

Diane R. Gold, MD, MPH, Suzanne Rogacz, MD, Naomi Bock, MD, MS, Tor D. Tosteson, DSc, Timothy M. Baum, MS, Frank E. Speizer, MD, and Charles A. Czeisler, PhD, MD

ABSTRACT

A hospital-based survey on shift work, sleep, and accidents was carried out among 635 Massachusetts nurses. In comparison to nurses who worked only day/evening shifts, rotators had more sleep/wake cycle disruption and nodded off more at work. Rotators had twice the odds of nodding off while driving to or from work and twice the odds of a reported accident or error related to sleepiness. Application of circadian principles to the design of hospital work schedules may result in improved health and safety for nurses and patients. (*Am J Public Health.* 1992;82:1011–1014)

Introduction

Concern for iatrogenic risks to patients caused by hospital staff experiencing sleep disruption has led to a reexamination of the work schedules and sleep patterns of physicians.[1–4] This cross-sectional study was designed to examine the impact of work schedule on the sleep schedule, sleepiness, and accident rates of female nurses in a Massachusetts hospital.

Methods

A self-administered questionnaire was distributed June through September 1986 to 878 registered nurses, licensed practical nurses, and other ancillary staff in the hospital. The nurse was asked to record, for the current week, the previous 2 weeks, and the following week, the number of shifts worked for each work-shift category (i.e., day, evening, night) at the hospital and at any other job. The nurse was also asked, "Does your job involve a variable work shift? That is, do you work the day shift sometimes and the night shift at other times?," which is a National Center for Health Statistics (NCHS) question previously used to identify shift workers in the US population.[5,6] In addition, the nurse recorded her sleep and wake times when she worked the day shift, the evening shift, the night shift, and on days off.

Information was collected regarding: quality of sleep; the use of alcohol (per month), prescription or nonprescription medication, sleeping aids, or other kinds of drugs to get to sleep; nodding off at work (per week); nodding off while driving to or from work in the past year; and accidents, errors, and "near-miss" accidents in the past year. The variable "any accident or error" included automobile accidents, medication errors, on-the-job procedural errors, and on-the-job personal injuries that the nurse reported had occurred because of sleepiness.

We used prior hypotheses based on physiologic and epidemiologic data regarding shift work and circadian disruption[7–17] in the definition of shift categories (Table 1), which were created from shift schedules for the current month. Day/evening shift work was grouped as a single category because shifting from days to evenings has not been demonstrated to disrupt circadian rhythms.

Sleep and wake times were used to determine whether a nurse obtained "anchor sleep": at least 4 hours of sleep obtained regularly during the same clock hours every night, both during work days and days off. Studies by Minors and Waterhouse[18] suggest that loss of anchor sleep may be a surrogate for circadian rhythm disruption.

The relationship between shift and outcomes such as sleepiness and accidents was explored first through univariate and then through multivariate analy-

Diane R. Gold, Naomi Bock, Tor D. Tosteson, Timothy M. Baum, and Frank E. Speizer are with the Channing Laboratory, and Suzanne Rogacz and Charles A. Czeisler are with the Center for Circadian and Sleep Disorders Medicine, Department of Medicine, Harvard Medical School, Brigham and Women's Hospital, Boston.

Requests for reprints should be sent to Diane R. Gold, MD, MPH, Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115.

This paper was submitted to the Journal March 6, 1991, and accepted with revisions November 12, 1991.

TABLE 1—Work Schedule Categories of Nurses Surveyed

Day/evening (n = 336): Within a month, working ≥ 4 day or evening shifts but no night shifts.
Night (n = 69): Within a month, working ≥ 8 night shifts and no day/evening shifts.
Rotator (n = 119): Within a month, working ≥ 4 day or evening shifts and ≥ 4 night shifts.
Day/evening, occasional night (n = 61): Within a month, working ≥ 4 day or evening shifts and 1–3 night shifts.
Night, occasional day/evening (n = 14): Within a month, working ≥ 8 night shifts and 1–3 day or evening shifts.
Part-time rotator (n = 17): Within a month, working 4–7 night shifts and 0–3 day or evening shifts.

ses, using EPISTAT and SAS (proc logistic). We attempted to validate the reporting of medication errors, but this was not possible because written accident reports submitted to the hospital's risk management office usually documented the nursing supervisor rather than the nurse associated with the error.

Results

Of 878 hospital employees contacted by the research assistants, 687 (78.3%) returned the questionnaire, 36 (4.1%) refused to participate, and 155 (17.7%) failed to return the questionnaire. The study cohort included 593 female registered nurses and 42 female licensed practical nurses.

The mean age of the nurses was 33.9 years (range 21 to 65 years). When compared to day/evening nurses, more rotators were 35 years old or younger (77.8% versus 64.5%). By contract, the option of not rotating and not working the night shift was dependent on seniority. Consequently, whereas 23.7% of rotators and 22.1% of night nurses had worked at the hospital for 1 year or less, only 9.3% of day/evening nurses had done so.

During their work days, rotators and night nurses reported fewer hours of sleep than day/evening nurses (Figure 1). Of the day/evening nurses, 92.2% obtained anchor sleep regularly throughout the month. In contrast, only 6.3% of night nurses, none of the rotators, and none of the nurses in work categories 4 through 6 (Table 1) obtained anchor sleep regularly throughout the month.

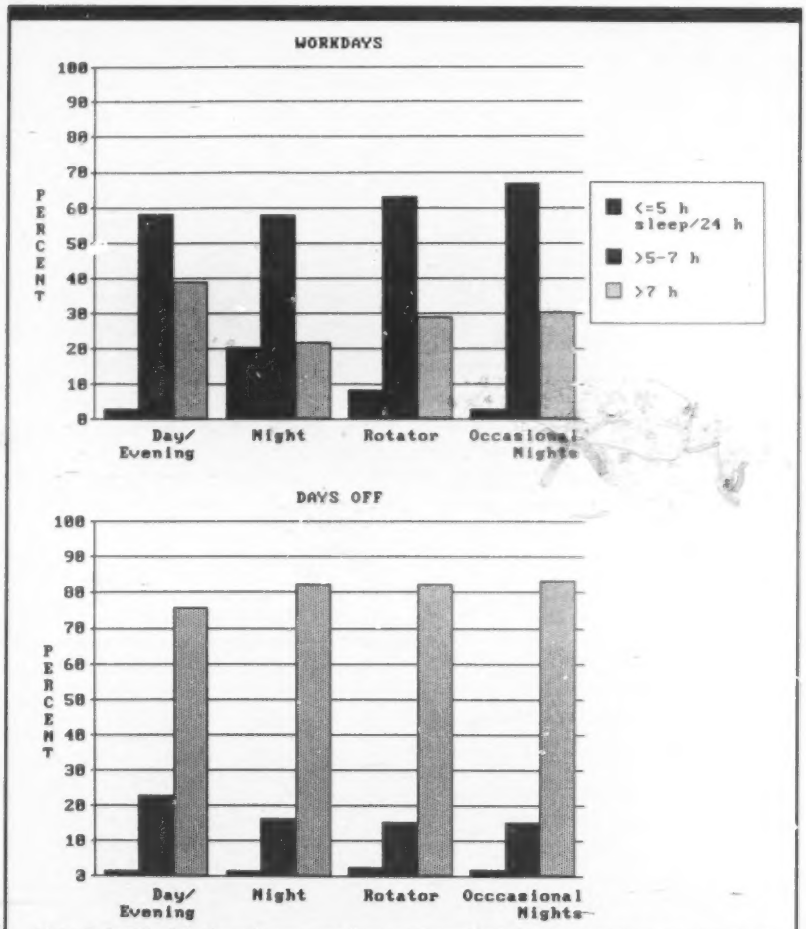
Of the nurses from all work categories who responded "yes" to the NCHS question on variable work schedules, only 53.6% experienced anchor sleep disruption. The group responding "yes" included 49.4% of the day/evening nurses, 94.1% of rotators, and 2.9% of night nurses.

In comparison to day/evening nurses, night workers had 1.8 times the odds and rotators had 2.8 times the odds of reporting poor quality sleep (Table 2). Night nurses and rotators had twice the odds of using medications to get to sleep.

Nodding off on the night shift occurred at least once per week in 35.3% of rotators, 32.4% of night nurses, and 20.7% of day/evening nurses who worked occasional nights. On the other hand, neither day/evening nurses nor rotating nurses reported significant problems with nodding off on the day or evening shift (rates for nodding off: 2.8% and 2.7%). When compared to day/evening nurses, rotators had 3.9 times the odds and night nurses had 3.6 times the odds of nodding off while driving to or from work in the preceding year (Table 2).

In univariate analyses two confounders of the relationship between shift and accidents were identified: working at the hospital 1 year or less was associated with medication errors, and an age of 35 years or younger predicted automobile near-miss accidents. The use of alcohol to get to sleep was an independent predictor of all categories of accidents.

Adjusting for these factors, the odds of reporting any accident or error were twice as high for rotators as for day/evening nurses (Table 3). Rotators had 2.5 times the odds of reporting near-miss accidents. After adjustment the effect of ro-



Note. Only 3% of day/evening nurses slept ≤ 5 hours/24 hours on workdays, as compared to 8% of rotators and 20% of night nurses. Less than 50% of nurses from each of the work schedules slept >7 hours/24 hours on workdays. On days-off greater than 75% of nurses from each of the work schedules slept >7 hours/24 hours.

FIGURE 1—Hours of sleep per 24 hours on workdays and on days off by category of work schedule.

TABLE 2—Differences between Categories of Work Schedules in Quality of Sleep, Sleepiness, and Use of Medications and Alcohol to Get to Sleep

	Day/Evening				Day/Evening, Occasional Night				Night				Rotator			
	n[a]	%	Odds Ratio[b]	95% CI	n[a]	%	Odds Ratio[b]	95% CI	n[a]	%	Odds Ratio[b]	95% CI	n[a]	%	Odds Ratio[b]	95% CI
Fatigue during the day	125	37.5	1.00	....	21	35.0	0.90	(0.48, 1.65)	19	27.9	0.65	(0.35, 1.19)	54	45.8	1.40	(0.90, 2.20)
Poor quality sleep	93	27.8	1.00	....	15	25.0	0.86	(0.44, 1.69)	28	41.2	1.81	(1.02, 3.22)*	62	52.1	2.82	(1.79, 4.45)*
Medications to get to sleep	55	16.5	1.00	....	15	25.0	1.69	(0.84, 3.39)	19	27.5	1.93	(1.01, 3.66)*	33	27.7	1.95	(1.15, 3.29)*
Alcohol to get to sleep	91	27.2	1.00	....	21	35.0	1.44	(0.77, 2.68)	17	24.6	0.88	(0.46, 1.65)	42	35.3	1.46	(0.91, 2.34)
Nodding off while driving to or from work	70	21.2	1.00	....	23	40.0	2.38	(1.27, 4.45)*	33	49.3	3.62	(2.02, 6.48)*	60	51.3	3.92	(2.45, 6.30)*

[a]Number of respondents answering "yes."
[b]Univariate unadjusted odds ratio.
*$P \leq 0.05$.

TABLE 3—Adjusted Odds Ratios[a] for Accidents/Errors among Nurses in Different Work Schedule Categories

	Day/Evening				Day/Evening, Occasional Night				Night				Rotator			
	n[b]	%	Odds Ratio[a]	95% CI	n[b]	%	Odds Ratio[a]	95% CI	n[b]	%	Odds Ratio[a]	95% CI	n[b]	%	Odds Ratio[a]	95% CI
Medication error	19	5.8	1.00	....	7	12.1	1.97	(0.76, 5.12)	5	7.4	1.17	(0.41, 3.34)	14	12.1	1.83	(0.86, 3.91)
Medication near miss	25	7.8	1.00	....	9	15.5	1.89	(0.82, 4.34)	11	16.4	2.10	(0.95, 4.68)	19	16.4	1.93	(0.99, 3.74)
Automobile accident	8	2.4	1.00	....	1	1.7	0.50	(0.06, 4.22)	3	4.4	2.24	(0.55, 9.07)	4	3.4	1.14	(0.33, 4.00)
Automobile near miss	65	19.6	1.00	....	16	28.1	1.47	(0.76, 2.83)	21	30.9	1.92*	(1.05, 3.52)	47	40.9	2.63*	(1.63, 4.24)
Any accident/error[c]	33	10.1	1.00	....	11	19.0	1.93	(0.88, 4.20)	11	16.2	1.88	(0.88, 4.02)	22	19.3	1.97*	(1.07, 3.63)
Any near miss[c]	86	26.8	1.00	....	22	38.6	1.57	(0.86, 2.86)	25	37.3	1.55	(0.87, 2.77)	58	50.4	2.47*	(1.56, 3.89)

[a]Odds ratio relative to day/evening workers; adjusted for age ≤ 35 yr, working at the hospital ≤ 1 yr, and the use of alcohol to get to sleep.
[b]Number of respondents answering "yes."
[c]Accidents/errors and near-miss accidents include automobile accidents, medication errors, on-the-job procedural errors, and on-the-job personal injuries that the nurses reported had occurred in the past year because of sleepiness.
*$P \leq 0.05$.

tating on medication errors was reduced from 2.2 (95% CI: 1.0, 4.9) to 1.8 (Table 3).

Discussion

The results of this study are consistent with laboratory investigations that have demonstrated that sleep deprivation and misalignment of circadian phase as experienced during rotating shift work are each associated with frequent lapses of attention and increased reaction time, leading to increased error rates on performance tasks.[7–9] These data also suggest that a record of a representative work schedule may be of greater use than the NCHS variable shift question in identifying workers with disrupted sleep/wake patterns.

Reporting bias and selection bias may have influenced the findings in this cross-sectional study. The hypotheses generated might be best tested through a case-control study design, with validation of reported accidents.

The data present a potential dilemma for hospital policymakers. Even if rotating leads to circadian rhythm disruption and accidents, nurses and hospital administrators will have to determine their priorities: the family responsibilities of nurses and the staffing requirements of hospitals may limit the potential for altering hospital schedules to improve the health of nurses and the safety of patients.[19,20] □

Acknowledgments

This research was supported in part by National Institutes of Health research and training grants NICHHD 5-R01-HD20174 and NHLBI 5-T32-HL07427 (Gold and Tosteson) and 5-MOI-RRO2635 (Rogacz).

We would like to acknowledge the enthusiastic participation of the hospital nurses, nursing administrators, and Department of Risk Management and Loss Control personnel in this research project. We are grateful to Mr. Joseph Ronda for assistance in data management and to Ms. Darlene Bramble for secretarial assistance.

References

1. Friedman RC, Bigger JT, Kornfield DS. The intern and sleep loss. *N Engl J Med.* 1971;285:201–203.
2. Lurie N, Rank B, Parenti C, Woolley T, Snoke W. How do house officers spend their nights? A time study of internal medicine house staff on call. *N Engl J Med.* 1989;320:1673–1677.
3. Orton DI, Gruzelier JH. Adverse changes in mood and cognitive performance of house officers after night duty. *Br Med J.*

1989;298:21–23.
4. Durnford S. Junior hospital doctors: tired and tested. *Br Med J.* 1988;297:931–932.
5. Danchik KM, Schoenborn CA, Elinson J. *Highlights from Wave 1 of the National Survey of Personal Health Practices and Consequences: United States.* Vital and health statistics; series 15, no. 1. Washington, DC: US Government Printing Office; 1981; US DHHS publication PHS 81-1162.
6. Gordon NP, Cleary PD, Parker CE, Czeisler CA. The prevalence and health impact of shiftwork. *Am J Public Health.* 1986;76:1225–1228.
7. Dinges DF, Orne MT, Whitehouse WG, Orne EC. Temporal placement of a nap for alertness: contributions of circadian phase and prior wakefulness. *Sleep.* 1987;10:313–329.
8. Hamilton P, Wilkinson RT, Edwards RS. A study of four days partial sleep deprivation. In: Colquhoun WP, ed. *Aspects of Human Efficiency.* London, England: English Universities Press, Ltd.; 1972:101–113.
9. Williams HL, Lubin A, Goodnow JJ. Impaired performance with acute sleep loss. *Psychological Monogr.* 1959;73:1–26.
10. Czeisler CA, Weitzman ED, Moore-Ede MC, Zimmerman JC, Knauer RS. Human sleep: its duration and organization depend on its circadian phase. *Science.* 1980;210:1264–1267.
11. Bjerner B, Swensson A. Schichtarbeit and rhythmus. *Acta Med Scand Suppl.* 1953;278:102–107.
12. Moore-Ede MC, Czeisler CA, Richardson GS. Circadian timekeeping in health and disease. *N Engl J Med.* 1983;309:530–536.
13. Akerstedt T. Sleepiness as a consequence of shift work. *Sleep.* 1988;11:17–34.
14. Folkard S, Monk TH, Lobban MC. Short and long-term adjustment of circadian rhythms in 'permanent' night nurses. *Ergonomics.* 1978;21:785–799.
15. Folkard S, Haines SM. Adjustment to night work in full and part-time night nurses. *J Physiol.* 1977;267:23P.
16. Adams J, Folkard S, Young M. Coping strategies used by nurses on night duty. *Ergonomics.* 1986;29:185–196.
17. Gadbois C. Women on night shift: interdependence of sleep and off-the-job activities. In: Reinberg A, Vieux N, Andauer P, eds. *Night and Shift Work Biological and Social Aspects.* New York, NY: Pergamon Press; 1981:223–227.
18. Minors DS, Waterhouse JM. Anchor sleep as a synchronizer of rhythms on abnormal routines. *Int J Chronobiology.* 1981;7:165–188.
19. Czeisler CA, Moore-Ede MC, Coleman RM. Rotating shift work schedules that disrupt sleep are improved by applying circadian principles. *Science.* 1982;217:460–463.
20. Knauth P, Rutenfranz J. Development of criteria for the design of shiftwork systems. *J Hum Ergol.* 1982;11:337–367.

Birthweight Distributions in Mexico City and among US Southwest Mexican Americans: The Effect of Altitude

Francis C. Notzon, PhD, José Luis Bobadilla, MD, PhD, and Irma Coria, MPH

ABSTRACT

In this study birthweight distributions for Mexican Americans and Mexico City were compared. Sharp differences in the two distributions were nearly eliminated by controlling for altitude. The small remaining excess in low birthweight in Mexico City appears to be due to a slight overrepresentation of tertiary hospital deliveries, and possibly to a greater prevalence of pathological conditions. The results are consistent with the favorable low birthweight rate reported for Mexican Americans and illustrate the need to adjust for altitude in studies of low birthweight. (*Am J Public Health.* 1992;82:1014–1017

Introduction

Low birthweight rates (birthweight <2500 g) vary widely with racial or ethnic origin.[1–3] However, these differences are attributed primarily to socioeconomic conditions strongly associated with these groups,[4–6] leading to the assumption that the more socioeconomically disadvantaged groups are the most likely to have low birthweight babies. Low socioeconomic status often is associated with factors known to impair the rate of fetal growth, such as inadequate nutrition and smoking.[7] Insufficient prenatal care and other factors also may mediate between low socioeconomic status and low birthweight.[4] While the association between socioeconomic status and ethnicity appears to hold for US Blacks and Puerto Ricans, this is not so for Mexican Americans.

The low birthweight rate for Mexican Americans is similar to that of US Whites, despite low socioeconomic status and poor access to prenatal care, and is far lower than the rate for US Blacks or Puerto Ricans (Table 1). This phenomenon, described as a "public health enigma," is not readily explainable.[8] Nevertheless, the low birthweight rate of Mexican Americans is so dramatically different from other US minority groups that a verification of the accuracy of the Mexican-American rate is warranted.

One approach to verifying these data is to compare Mexican-American birthweights with those found in Mexico. If

Francis C. Notzon is with the National Center for Health Statistics, Centers for Disease Control, US Department of Health and Human Services, Hyattsville, Md. José Luis Bobadilla is with the Population, Health and Nutrition Division, The World Bank, Washington, DC. Irma Coria is with the Department of Analysis and Statistics, National Institute of Perinatology, Mexico City, DF, Mexico.

Requests for reprints should be sent to Francis C. Notzon, PhD, National Center for Health Statistics, 6525 Belcrest Road, Hyattsville, MD 20782.

This paper was submitted to the Journal December 4, 1990, and accepted with revisions November 27, 1991.

TABLE 1—Low Birthweight Rate by Hispanic Origin and Race, 23 Reporting US States and the District of Columbia, 1984

Origin/Race	Births under 2500 g, %
Hispanic	
Mexican	5.7
Cuban	5.9
Puerto Rican	8.9
Non-Hispanic	
White	5.5
Black	12.4

Source: National Center for Health Statistics, Ventura SJ.[2]

TABLE 2—Low Birthweight Rates among Live Births, US Southwestern Mexican Americans and Mexico City, 1984

	Births under 2500 g, %	Median Birthweight, g
US Southwest Mexican Americans	5.9	3374
Mexico City	10.6	3100

Sources: US National Center for Health Statistics, national vital statistics data, 1984; Mexico City, Center for Public Health Research, perinatal mortality study, 1984.

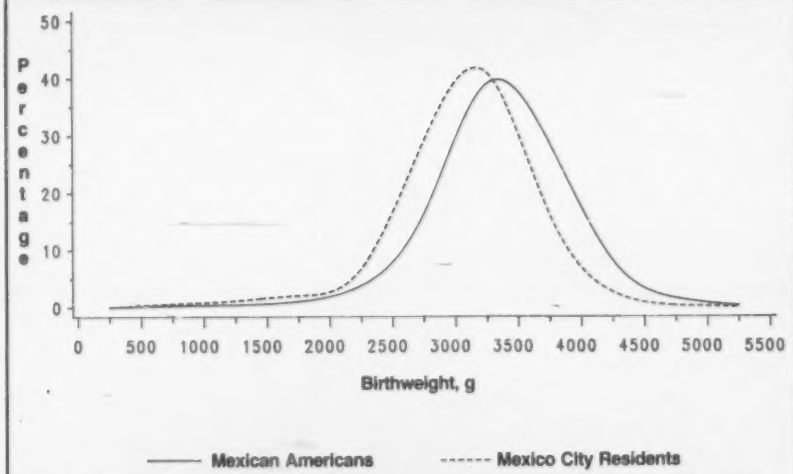


FIGURE 1—Birthweight distributions: US Southwest Mexican Americans, 1980–1984, and Mexico City Residents, 1984. Sources: US National Center for Health Statistics, national vital statistics data, 1980–1984; Mexico City, Center for Public Health Research, perinatal mortality study, 1984.

Mexican-American birthweights truly are not strongly affected by adverse health care and economic circumstances, then one would expect to find a roughly similar birthweight distribution in Mexico, at least in the more economically advanced areas. Genetic differences between Mexican Americans and Mexicans should be small, and behavioral factors such as low usage levels of tobacco, alcohol, and illegal drugs are common to both groups and likely to be lower for Mexican women.[9]

While accurate and complete information on birthweight is not available for the total population of Mexico, a 1984 study collected such data on births in Mexico City.[10] However, inhabitants of Mexico City reside at an altitude of 2200 m (about 7200 ft), while US Southwest Mexican-American births occur primarily at elevations below 500 m (about 1600 ft).

This study compares birthweight distributions for these two groups, using information on altitude of county of residence to adjust for differences in elevation between the two populations.

Materials and Methods

Information on Mexican-American births was obtained from US vital statistics for five southwestern states (Arizona, California, Colorado, New Mexico, Texas) for the years 1980 through 1984. These states contain 92% of all Mexican-American live births registered in the United States.[2] Vital statistics for the years 1980 through 1984 were combined into a single file, including about 1.1 million live births to women of Mexican-American or other Spanish origin in these states.

Information on the mean altitude of each county in the five-state area was obtained from the Bureau of Health Professions Area Resources File.[11] Sufficient numbers of Mexican-American births were available to permit determination of birthweight measures at 500 m gradations, up to the range of 2000 to 2499 m.

Information on births in Mexico City came from a 1984 study of the quality of medical care and perinatal mortality in Mexico City hospitals conducted by the Mexican Ministry of Health.[10] The study collected information on 32 760 live births occurring during May through August 1984 in 25 hospitals in the Federal District of Mexico (37% of all hospital births in Mexico City). Extensive quality-control measures were undertaken to ensure that birthweights were accurately recorded for all births.

Sample hospitals were selected from all Mexico City hospitals, stratified by health care agency (public assistance, social security, and private) and, for social security hospitals, by level of care. The resulting sample can be considered as approximately representative of all hospital births in Mexico City, although tertiary hospital births are slightly overrepresented.

The overwhelming majority of births in Mexico City in the 1980s occurred in hospitals. The National Fertility and Health Survey (1987) found that about 97% of Mexico City births between 1982 and 1986 were delivered by a medical doctor, and almost all occurred in hospitals.[12]

Birthweight outcomes for births in Mexico City and for Mexican Americans from southwestern states were compared through the use of various birthweight measures (medians, percentage under 1500 g, percentage under 2500 g) and by graphing altitude-specific birthweight distributions for the two groups. Graphs of the various birthweight distributions were produced using the SAS/GRAPH package.[13] The curves were based on 500-g intervals, smoothed using cubic spline interpolation.

Results

A comparison of birthweights in the two groups without adjustment for altitude shows a sharply higher low birthweight rate in Mexico City, almost double the Mexican-American rate, and a differ-

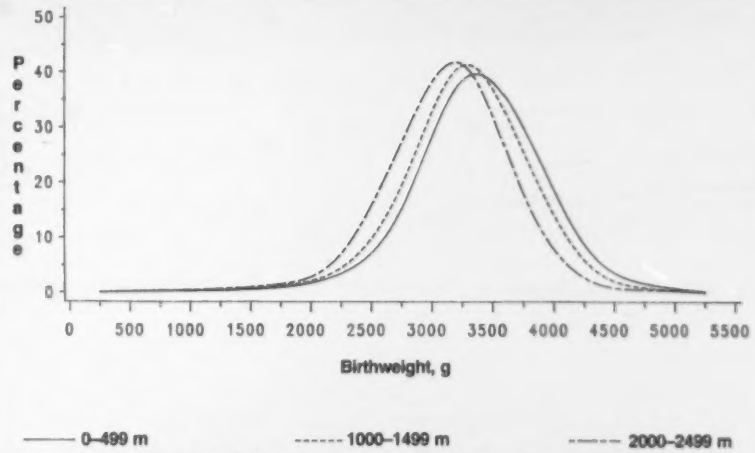


FIGURE 2—Birthweight distributions by altitude: US Southwest Mexican Americans, 1980 to 1984. Sources: US National Center for Health Statistics, national vital statistics data, 1980 to 1984, and Bureau of Health Professions, Area Resources File, 1980.

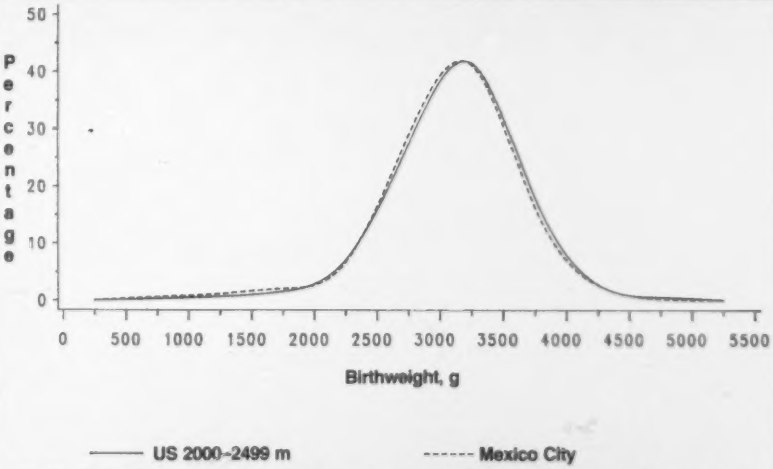


FIGURE 3—Birthweight distributions, adjusted for altitude: US Southwest Mexican Americans, 1980 to 1984, and Mexico City Residents, 1984. Sources: US National Center for Health Statistics, national vital statistics data, 1980 to 1984, and Bureau of Health Professions, Area Resources File, 1980; Mexico City, Center for Public Health Research, perinatal mortality study, 1984.

ence in median birthweights of almost 275 g (Table 2). Comparing the percentage distributions of birthweights in the two populations shows that strong differences exist throughout the range of birthweights (Figure 1).

Differences in altitude have been shown to have a strong impact on birthweights for US Whites.[14,15] Altitude is thought to retard fetal growth because of the lower oxygen content of air at higher elevations, resulting in fetal hypoxia. The result is to increase not only the proportion of births below 2500 g, but to reduce birthweights throughout the distribution.

Figure 2 provides strong evidence that Mexican-American birthweight also is affected by altitude. As shown in the figure, the Mexican-American distribution shifts leftward with each increase in altitude. Figure 3 compares the Mexico City distribution with the Mexican-American curve for the most similar altitude, that is, for births to mothers residing at 2000 to 2499 m. The two curves are quite similar, and for births weighing more than 2500 g they virtually overlay. Differences remain for low birthweights, but much reduced from the curves presented in Figure 1.

Table 3 shows that the association between altitude and birthweight is present even at relatively low levels of elevation, as median birthweights drop consistently with each 500-m increment in altitude (χ^2 test for trend, $P < .001$). This table also shows that by controlling for altitude, the difference in low birthweight rates has been sharply reduced, so that the Mexico City rate is less than 10% greater than the Mexican-American rate at the same altitude (10.6% vs 9.7%, respectively). In addition, the difference in median birthweight has been reduced from 274 g to 47 g.

An alternative explanation for the association between altitude and birthweight is that altitude and socioeconomic status could be negatively correlated in the United States, so that poorer Mexican Americans would tend to live at higher altitudes. However, controlling for maternal education, as a proxy for socioeconomic status, has little effect on either the Mexico City/Mexican-American comparison or on the strength or consistency of the effect of altitude on birthweight (data not shown).

Discussion

The results presented in Figure 2 and Table 2 provide compelling evidence of the strength of the effect of altitude on birthweight. An increase in altitude of as little as 500 m affects not only the proportion of births under 2500 g, but causes the entire birthweight distribution to shift leftward. An increase of 2000 m almost doubles the percentage of low birthweight among Mexican Americans (from 5.5% to 9.7%), while the median birthweight declines by 227 g (3374 g to 3147 g). These results illustrate how misleading low birthweight rates, unadjusted for altitude, can be as indicators of health status or of general economic development.

The larger proportion of births under 2500 g in Mexico City, as compared with those of Mexican Americans, underlines the greater problem with pathological conditions in the Federal District, and presumably elsewhere in Mexico. A part of this excess however, may be due to the overrepresentation of tertiary-level hospitals, and thus of high-risk deliveries, in the Mexico City study. In addition, some differences in nutritional status remain between the two groups: maternal weight is almost 9 kg lower for Mexico City mothers (53.5 kg and 62.4 kg, respectively), and there is a 5-cm difference in their average

TABLE 3—Rates of Very Low Birthweight and Low Birthweight, and Median Birthweights, by Altitude: US Southwest Mexican Americans, 1980 to 1984, and Mexico City, 1984

Population/Altitude (m)	Births under 1500 g, %	Births under 2500 g, %	Median Birthweight, g	n
US Mexican Americans				
0–499	0.9	5.5	3360	907,965
500–999	1.0	6.8	3317	52,769
1000–1499	1.0	6.8	3289	79,934
1500–1999	1.0	8.5	3204	49,262
2000–2499	1.2	9.7	3147	11,798
Mexico City				
2200	1.8	10.6	3100	32,760

Sources: US National Center for Health Statistics, national vital statistics data, 1980–1984, and Bureau of Health Professions, Area Resources File, 1980; Mexico City, Center of Public Health Research, perinatal mortality study, 1984.

height (152.9 cm for Mexico City mothers vs 157.7 cm for Mexican Americans).[16,17]

By demonstrating the similarity in birthweight distributions between the Mexico City population and Mexican Americans, once altitude has been taken into account, these results help to confirm the validity of Mexican-American birthweight data. In addition, they suggest that adverse living conditions are not major determinants of birthweight in either of the two populations. The results also support previous research indicating that the lifestyle of Mexican-American women provides protection against the effect of adverse socioeconomic conditions on birthweight, including low smoking rates and the presence of social networks.[18] In addition, Mexican Americans have a higher maternal ponderal index and a lower proportion of primiparous births as compared with White non-Hispanics, which can also improve the birthweight distribution.[19] Further clarification of this "public health enigma" will require additional research. □

References

1. Evans S, Alberman E, Pashley J, Hampton B. International Collaborative Effort (ICE) on birthweight; plurality; and perinatal and infant mortality. II: Comparisons between birthweight distributions of births in member countries from 1970 to 1984. *Acta Obstet Gynecol.* 1989;68:11–17.
2. National Center for Health Statistics, Ventura SJ. *Births of Hispanic Parentage 1983 and 1984.* Hyattsville, Md: Public Health Service; 1987. DHHS publication no. 87-1120.
3. Shiono P, Klebanoff MA, Graubard BI, Berendes HW, Rhodes GG. Birth weight among women of different ethnic groups. *JAMA.* 1986;255:48–52.
4. Kleinman JC, Madans JH. The effects of maternal smoking, physical stature and educational attainment on the incidence of low birth weight. *Am J Epidemiol.* 1985; 121:843–855.
5. Institute of Medicine. *Preventing Low Birthweight.* Washington, DC: National Academy Press; 1985.
6. Victora CG, Barros FC, Vaughan JP, Teixeira AM. Birthweight and infant mortality: a longitudinal study of 5194 Brazilian children. *Int J Epidemiol.* 1987;16:239–245.
7. US Department of Health and Human Services. *Report of the Secretary's Task Force on Black and Minority Health. Volume VI: Infant Mortality and Low Birthweight.* Washington, DC: US Department of Health and Human Services; 1986.
8. Williams RL, Binkin NJ, Clingham EJ. Pregnancy outcomes among Spanish-surname women in California. *Am J Public Health.* 1986;76:387–391.
9. Hendershot GE, Zarate AO. The health of the foreign born population of the United States: data from the National Health Interview Survey, 1985. Presented at the annual meeting of the Population Association of America; April 29–May 2, 1987; Chicago, Ill.
10. Bobadilla JL. *Quality of Perinatal Medical Care in Mexico City.* Mexico City: Secretaría de Salud, Instituto Nacional de Salud Pública; 1988.
11. Bureau of Health Professions. *Area Resources File System.* Rockville, Md: Public Health Service; 1989.
12. Secretaría de Salud (México). *Encuesta Nacional sobre Fecundidad y Salud 1987.* Mexico City, DF, Mexico: Secretaría de Salud and Institute for Resource Development; 1989.
13. SAS Institute Inc. *SAS/GRAPH User's Guide, Version 5 Edition.* Carey, NC: SAS Institute Inc; 1985.
14. Yip R. Altitude and birth weight. *J Pediatr.* 1987;111:869–876.
15. Unger C, Weiser JK, McCullough RE, Keefer S, Moore LG. Altitude, low birth weight, and infant mortality in Colorado. *JAMA.* 1988;259:3427–3432.
16. Coria I, Bobadilla JL, Cerón S. Valores de referencia para evaluar el crecimiento intrauterino en nacimientos occuridos en la Ciudad de México. *Salud Pública de México.* 1988;30:68–80.
17. Najjar MF, Kuczmarski RJ. *Anthropometric Data and Prevalence of Overweight for Hispanics: 1982–84.* Hyattsville, Md: National Center for Health Statistics; 1989. (Vital and health statistics; series II; no. 239).
18. Scribner R, Dwyer JH. Acculturation and low birthweight among Latinos in the Hispanic HANES. *Am J Public Health.* 1989; 79:1263–1267.
19. Notzon FC. *Factors Associated with Low Birth Weight in Mexican Americans and Mexicans.* Baltimore, Md: Johns Hopkins; 1989. Thesis.

Identification of Nonfatal Family and Intimate Assault Incidents in Police Data

Linda E. Saltzman, PhD, James A. Mercy, PhD, and Philip H. Rhodes, MS

ABSTRACT

We examined different strategies for identifying nonfatal family and intimate assaults (FIAs) in police data. Police most often classify such incidents in the "assault" category, but they also use other crime categories. We estimated that, during 1984, 3300 FIAs (or 837 per 100 000 population) occurred in Atlanta. Of those, 77% were classified as assaults; 23% were classified in nonassault categories. Research measuring the magnitude of FIAs should take into account incidents classified in nonassault crime categories. (*Am J Public Health.* 1992;82:1018–1020)

Introduction

Police data are a systematic, ongoing collection of information about crime and a valuable resource in examining the epidemiological patterns and causes of violence among intimates.[1,2] Police data on family and intimate assaults (FIAs) can benefit communities in several ways: a minimum FIA rate can be estimated; groups at high risk of FIA involvement can be identified to target prevention strategies; and the data can be used to help evaluate intervention strategy effectiveness.

We describe how one city's police incident reports were sampled to estimate the number of nonfatal FIAs in a year, and how different strategies for identifying FIA incidents affect rate estimates.

Methods

As part of a larger study,[3] we reviewed nonfatal incident reports for 1984 from the Bureau of Police Services in Atlanta, Ga. FIAs met four criteria: (1) incident reports described one or more of the following circumstances: physical contact between parties, use of a weapon, threat with a weapon, or verbal threat of bodily harm; (2) the incident had occurred in the city of Atlanta and within Fulton County; (3) the victim's residence was Atlanta within Fulton County; and (4) the victim and perpetrator were emotionally intimate at the time of the incident or before it occurred.

Threats and physical contact causing no discernible injuries were included as FIAs because they often escalate to violence with serious injury or homicide.[4] We required both incident location and victim's residence to be within Atlanta and Fulton County because (1) such victims were eligible for county agency services and the larger study required records reviews at public agencies; and (2) population estimates were available for that defined geographic area, enabling calculation of FIA rates.

Incidents were eligible for classification as FIAs if the perpetrator-to-victim relationship was included among a list of relationships implying closeness or emotional intimacy (e.g., a married couple). The list included familial relationships, former relationships, and relationships without legal definition (e.g., a cohabiting or dating couple). For incidents with multiple perpetrators, at least one perpetrator's relationship to the victim had to be emotionally intimate. Incident reports without relationship information (nearly 25% of reports reviewed) were considered ineligible.

Atlanta police incident reports are filed within Federal Bureau of Investigation (FBI) Uniform Crime Report[5] (UCR) categories. In keeping with the UCR's crime hierarchy, assaults that occur during incidents of homicide, rape, and robbery are classified in those "higher-order" crime categories and not as assaults. Thus, reviewing cases from the assault crime category automatically misses some incidents involving assault. Similarly, several "lower-order" categories include incidents meeting our case definition.

Because the proportion of FIAs classified in any of these nonassault crime categories was unknown, we did a pilot review. We took a simple random sample of 1984 incident reports in each of seven nonfatal FBI crime categories. The sample sizes in the strata ranged from 1% (assault, a large category with 10 560 incident reports) to 15% (family and children offenses, a small category with only 192 incident reports). The pilot sample provided a working guess as to the relative number of records contributed by each of the seven crime categories. This initial assessment was done to establish guidelines for selecting the actual study sample, not to provide final estimates of the numbers of FIAs.

The authors are with the Centers for Disease Control's National Center for Environmental Health and Injury Control in Atlanta, GA.

Requests for reprints should be sent to Linda E. Saltzman, Centers for Disease Control, 1600 Clifton Road, MS-F-36, Atlanta, GA 30333.

This paper was submitted to the Journal March 6, 1991, and accepted with revisions October 18, 1991.

Using the pilot results, we then set target numbers for FIAs within each strata (total n across all strata = 150). Within each crime category, we randomly selected and reviewed records until we obtained the target number of FIAs (i.e., negative binomial sampling). We reviewed a total of 945 incident reports in the seven nonfatal crime categories (519 classified as assault and 426 incident reports classified in the other six categories) to obtain 150 cases satisfying the case definition for FIA. Case distribution reflected the estimated percent of nonfatal FIAs in 1984 contributed by each crime classification: rape (n = 2); robbery (n = 3); assault (n = 125); disorderly conduct (n = 13); family and children offense (n = 2); sex offense (n = 1); and other (n = 4).

To estimate FIA incidence, we multiplied the known number of incident reports in each of the seven nonfatal crime categories by the estimated proportion of records within each crime category meeting the FIA case definition. The total estimated number of FIAs is the sum of the estimates for the seven crime categories. The incidence estimates were then used to calculate 1984 nonfatal FIA victimization rates for (1) incidents classified in all seven nonfatal crime categories, (2) incidents classified only in the assault crime category, and (3) incidents classified as nonassault crimes. Population estimates for the city of Atlanta within Fulton County came from the Atlanta Regional Commission's 1985 population update,[6] because complete population data were not available for 1984.

Approximate 95% confidence intervals (CI) for the numbers and rates per 100 000 population were calculated based on the actual sampling design, but other calculations indicated that an assumption of simple stratified random sampling within each crime category would have given almost identical results. Confidence interval calculations are based on an assumption of a normal distribution. The distribution of the log (number of events) meets this assumption better than the distribution of the number of events. Therefore, the confidence intervals are not symmetric around the point estimates because they were initially calculated around the log (number of events) and then converted back to the original scale. Within each crime category, the probabilities are closer to 0 than to .5 that any particular police incident report meets the FIA case definition. Thus, the confidence intervals extend slightly higher above the point estimates than below them.

TABLE 1—Incidence Estimates, Nonfatal Family and Intimate Assaults (FIAs), Atlanta, Georgia, 1984

Incident Report Classification	Number of Incident Reports	Estimated FIAs (95% CI)	Percent of Total FIAs	Percent FIAs in Crime Category
Assault	10 560	2541 (2175, 2959)	77	24
Nonassault Crime[a]	12 996	759 (513, 1112)	23	6
Total	23 556	3300 (2846, 3821)	100	14

[a] Nonassault crime classifications include rape, robbery, disorderly conduct, family and children offense, sex offense, and other.

Results

We estimated that 3300 (14%) of the 23 556 incident reports filed in the seven crime categories in Atlanta in 1984 were FIAs (Table 1). Of the 10 560 incidents classified as assault, we estimated that 24% were FIAs; and of the 12 996 incident reports in the other six categories, we estimated that 6% were FIAs. Of the estimated 3300 FIAs, 77% were contributed by incidents from the assault category and 23% were contributed by incidents from the combined nonassault crime categories.

The FIA rate based on all seven nonfatal crime categories was 837 per 100 000 population (95% CI = 724–967). The FIA rate based only on incidents classified as assault was 645 per 100 000 population (95% CI = 552–749).

Discussion

Although we estimated that 77% of FIA incidents in Atlanta in 1984 were classified as crimes of assault, the number of FIAs classified as nonassault crimes was not negligible. Our finding that some FIA incidents were classified in higher-order UCR crime categories (rape, robbery) is consistent with analysis of National Crime Survey data; of reported violent incidents among intimates, 89% involved the crime of assault, 9% involved robbery, and 2% involved rape.[7]

The 30% increase in the FIA victimization rate when nonassault-classified FIAs are combined with assault-classified FIAs suggests that analyses using police data should either include nonassault-classified FIAs or acknowledge the likelihood that estimates solely from assault-classified FIAs are underestimates. Much intimate violence never comes to the attention of the police,[8,9] so even with the addition of nonassault-classified FIAs, police data will underestimate FIA incidence and victimization rates. For many communities, however, police data provide the *only* routinely collected information about FIAs. Furthermore, the police are often the first to contact assault participants, and police can provide information about those encounters that is unavailable from other sources.

Because UCR data collection has not routinely included information about the perpetrator-to-victim relationship for nonfatal incidents, it was not surprising that nearly a fourth of the incident reports we reviewed lacked such information. Imminent changes occurring with the FBI's National Incident-Based Reporting System (NIBRS) should greatly improve the utility of police data for studying FIAs because NIBRS will routinely include perpetrator-to-victim relationship information for nonfatal crimes.[10]

Communities wishing to establish FIA surveillance might consider the following recommendations.

1. Determine what police data on FIA are already available in the community, and explore the feasibility of undertaking new or expanded data collection. In particular, consider expanding routinely recorded information to include the perpetrator-to-victim relationship.

2. In large jurisdictions, sample police incident report records, recognizing that the resources needed to collect information on all FIA incidents would be great. Smaller jurisdictions may be able to collect information about all incidents more easily.

3. Incorporate not only incidents classified in additional nonfatal crime cat-

egories besides assault but also homicide incidents.

4. Use any resulting data to modify or initiate community policies that address prevention of family violence and other assaults among intimates. It will help if efforts can be made to establish (or strengthen) communication links between criminal justice and public health agencies, and between those agencies and community advocacy groups concerned about FIA. □

Acknowledgments

This paper uses data from the study of Family and Intimate Assaults in Atlanta (FINAA), supported under a collaborative agreement (#U50/CCU400931-01-1) between the Centers for Disease Control and the Georgia Department of Human Resources, with the additional collaboration of the Fulton County (Georgia) Health Department and the Atlanta Department of Public Safety.* Additional support was provided through an interagency agreement (#86-IJ-R-200) with the National Institute of Justice.

This paper was presented in November 1988 at the annual meeting of the American Society of Criminology in Chicago, Ill.

We thank Robert Nash Parker for his helpful comments on a previous draft of this article.

*The following were members of the Collaborative Working Group for the FINAA study: (1) Atlanta Department of Public Safety: George Napper, PhD; Ardith Peters, PhD; Guy O. Seymour, PhD; and Jack Mallory, MSW; (2) Fulton County Health Department: William R. Elsea, MD, MPH; Robert J. Fiston, MSPH; and Richard F. Lyles, PhD; (3) Georgia Department of Human Resources: Project Director R. Keith Sikes, DVM, MPH; Assistant Project Director Jane C. Carr, RN; Senior Research Associate Sandra S. Huguley, MA; Research Associates Anne Becker, Lois Ellison, Steve Erickson, Theresa Irby, Debbie Lester, Kathy Marth, William D. Patterson, MA, Terrie Sterling, MA, MS, and Angela Wine; (4) Centers for Disease Control: Project Officer Linda E. Saltzman, PhD; James A. Mercy, PhD; Mark L. Rosenberg, MD, MPP; Richard J. Waxweiler, PhD; and Harvey F. Davis, Jr, MPH.

References

1. Weis JG. Family violence research methodology and design. In: Ohlin L, Tonry M, eds. *Family Violence*. Chicago, Ill: University of Chicago Press;1989:117–162.
2. Wilt GM, Bannon JD, Breedlove RK, et al. *Domestic Violence and the Police: Studies in Detroit and Kansas City*. Washington, DC: The Police Foundation;1977.
3. Saltzman LE, Mercy JA, Rosenberg ML, et al. Magnitude and patterns of family and intimate assault in Atlanta. *Violence and Victims*. 1990;5:3–17.
4. Straus MA: Domestic violence and homicide antecedents. *Bull N Y Acad Med*. 1986;62:446–465.
5. US Department of Justice. *Crime in the United States, 1988*. Washington, DC: US Govt Printing Office; 1989.
6. Atlanta Regional Commission. *Update '85: Age, Race, Sex*. Atlanta, Ga: Atlanta Regional Commission; 1986.
7. US Department of Justice. *Intimate Victims: A Study of Violence among Friends and Relatives*. Washington, DC: US Govt Printing Office; 1980.
8. Schulman MA. *A Survey of Spousal Violence against Women in Kentucky*. Washington, DC: US Dept of Justice; 1979.
9. Teske RHC Jr, Parker ML. *Spouse Abuse in Texas: A Study of Women's Attitudes and Experiences*. Huntsville, Tex: Survey Research Program, Criminal Justice Center, Sam Houston State University; 1983.
10. US Department of Justice. *Uniform Crime Reporting: National Incident-Based Reporting System*. Vol 1: *Data Collection Guidelines*. Washington, DC: US Govt Printing Office; 1988.

The Cost and Frequency of Hospitalization for Fall-Related Injuries in Older Adults

Bruce H. Alexander, MS, Frederick P. Rivara, MD, MPH, and Marsha E. Wolf, PhD

ABSTRACT

Using a population-based hospital discharge registry with E codes, we examine the 1989 hospitalizations of older adults in Washington State for fall-related injuries. Fall-related trauma accounted for 5.3% of all hospitalizations of older adults, with hospital charges totaling $53 346 191, and resulted in discharge to nursing care more often than other such hospitalizations. An annual hospitalization rate of 13.5 per 1000 persons and an annual cost of $92 per person is reported. The importance of preventing fall-related injuries in older adults is discussed. (*Am J Public Health*. 1992;82:1020–1023)

Introduction

The risk of falls and of fall-related injury increases with age owing to loss of agility and visual acuity, predisposition to dizziness and syncope, and side effects from medications.[1–7] Approximately 30% of older adults will experience a fall annually, and 20% to 30% of those people will suffer moderate to severe injuries[6,8] leading to loss of mobility and independence and to an increased risk of death.[1,2,9–11] A growing body of literature is available on risk factors for fall-related injuries, especially fracture of the hip, and has been reviewed by others elsewhere.[1,2,5] However, complete data on these injuries in a population, along with attendant cost and morbidity information, have not been well presented. To characterize the frequency and cost of fall-related injuries that result in hospitalization, a population-based study was conducted of all persons aged 65 years and older in Washington State who were hospitalized in 1989 after a fall.

Methods

Hospital discharge data were obtained from the 1989 Washington State

All three authors are with the Harborview Injury Prevention and Research Center, and the University of Washington's Department of Epidemiology in Seattle; Frederick P. Rivara is also with the University of Washington's Department of Pediatrics.

Requests for reprints should be sent to Dr. Frederick P. Rivara, Harborview Injury Prevention and Research Center, Mailstop ZX-10, 325 Ninth Avenue, Seattle, WA 98104.

This paper was submitted to the Journal September 3, 1991, and accepted with revisions December 31, 1991.

TABLE 1—Characteristics of Patients by Discharge Category

Variable	Trauma			Nontrauma			Total (n = 149 504) %
	Fall Related (n = 7873) %	Nonfall Related (n = 1371) %	Cause Unknown (n = 2357) %	Noninjury (n = 123 759) %	Injury, Cause Known (n = 2066) %	Injury, Cause Unknown (n = 12 078) %	
Age, y							
65–69	12.4	26.7	19.9	23.9	24.6	29.0	23.7
70–74	15.5	24.1	21.1	24.3	25.6	27.9	24.1
75–79	19.1	20.0	21.0	21.9	21.2	22.1	21.7
80–84	22.0	16.0	17.8	16.0	16.7	13.0	16.2
85+	31.1	13.3	20.3	13.8	11.9	7.9	14.3
Gender							
Female	73.6	51.2	58.8	53.5	55.7	46.5	54.1
Male	26.4	48.8	41.2	46.5	44.3	53.5	45.9
Discharge status							
Self care	42.7	68.6	60.9	72.4	71.5	73.4	70.7
Nursing care	42.8	16.6	21.5	11.8	14.0	10.3	13.5
Hospital	2.5	3.0	2.9	3.0	3.3	2.1	2.9
Home/aide	8.9	7.7	9.0	6.7	6.9	8.3	7.0
Left against medical advice	0.1	0.1	0.1	0.2	0.4	0.1	0.2
Died	2.9	4.2	5.6	6.0	3.9	5.8	5.8
Hip fracture[b]	55.2	16.0	23.9	...	...	...	3.4
Other fracture	28.6	39.2	20.2	...	...	...	2.2

Note. Totals may fall short of or exceed 100% because of rounding.
[a]*ICD* discharge codes 800–904, 910–957.
[b]Fracture of the hip, femur, or pelvis (*ICD* 820, 821, 808).

Commission Hospital Abstract Reporting System, which collects discharge data from all Washington State hospitals, excluding Veterans Administration and military hospitals.

Patients were initially divided into the following six mutually exclusive categories, based on *International Classification of Diseases* (*ICD*) discharge codes and E codes indicating, respectively, presence or absence of trauma and the external cause of injury.

1. *Fall-related trauma: ICD* trauma code (800–904, 910–957) and an E code for a fall (E880–E889).

2. *Nonfall-related trauma: ICD* trauma code and an E code for trauma excluding falls and intentional injury (E800–E848, E890–E899, E916–E928).

3. *Trauma of unknown cause: ICD* trauma code but no E code recorded.

4. *Noninjury hospitalizations:* All *ICD* discharge codes lower than 800.

5. *Nontrauma injuries: ICD* code for late effects from external cause (905–909) or late effects of trauma, poisonings, or medical misadventure (958–998); and an E code for poisoning (E850–E869), medical misadventure (E870–E876), environmental effects (E900–E909), suffocation or submersion (E910–E915), late effects of injury (E929), adverse effects of medication (E930–E949), or intentional injury (E950–E999).

6. *Nontrauma injuries, unknown cause: ICD* codes 905–909, 958–998 but no E code recorded. Groups 4, 5, and 6 were ultimately combined to form an "all nontrauma" category. Population-based rates were calculated using age- and sex-specific denominator data for Washington in 1989 based on projections from 1980 and 1990 census data.[12]

The frequency of hospitalizations in each discharge category for persons over age 65 in Washington State in 1989 was broken down by age, gender, source of admission, and discharge status; and age- and sex-specific hospitalization rates for falls, other trauma, and nontrauma among this population were calculated. The cost of fall-related injuries was based on hospital charges only, exclusive of professional fees.

Results

Of the 149 504 patients aged 65 years and older who were discharged from Washington State hospitals in 1989, 7873 (5.3%) were hospitalized for injuries from falls (Table 1). Compared with persons hospitalized for other reasons, persons admitted for falls were more frequently female and tended to be older.

Patients hospitalized for fall-related trauma were discharged to a nursing facility more often than were those hospitalized for other reasons (Table 1). This finding was independent of age and patient status prior to admission. Patients admitted from home with fall-related injuries were discharged to a nursing facility more often than were patients hospitalized both for other trauma combined and for reasons not related to trauma (42% vs 23% and 14%, respectively) (Table 2).

The rate of hospitalization for fall-related trauma, 13.5 per 1000 persons aged 65 and older, was five times that of nonfall-related trauma (Table 3). The rate for falls increased with age and was higher among women than among men.

Of the $995 499 233 in hospital charges in 1989 for persons in Washington State aged 65 and older, $53 346 191 (5.3%) were attributable to hospitalization of patients with fall-related trauma. The annual per capita cost of such hospitalizations was higher for women and increased remarkably with age (Table 4).

Discussion

Population-based hospital discharge data with E codes were used to describe

TABLE 2—Age-Adjusted Distribution of Discharge Destinations for Patients Admitted from Self Care

Destination upon Discharge	Trauma[a] Fall Related (n = 7485) %	Trauma[a] Nonfall Related (n = 1296) %	Trauma[a] Unknown Cause (n = 2180) %	Nontrauma[b] (n = 137 903) %
Self care	43.3	62.8	58.0	68.5
Nursing care	42.1	21.4	23.4	14.1
Hospital	2.5	2.4	2.8	2.1
Home/aide	9.1	9.1	9.5	6.9
Left against medical advice	0.2	<0.1	<0.1	0.1
Died	2.9	5.9	5.9	6.3

Note. Totals may fall short of or exceed 100% because of rounding.
[a]*ICD* discharge codes 800–904, 910–957.
[b]All nontrauma combined.

TABLE 3—Age- and Sex-Specific Rates of Hospitalization per 1000 Adults Aged 65 Years and Older in Washington State in 1989

Sex and Age, y	Trauma[a] Fall Related (n = 7873)	Trauma[a] Nonfall Related (n = 1371)	Trauma[a] Unknown Cause (n =2357)	Nontrauma[b] (n = 137 903)	Total (n = 149 504)
Female	17.3	2.1	4.2	218.5	242.1
65–69	6.2	1.6	2.2	158.3	168.2
70–74	10.0	2.0	3.1	202.3	217.4
75–79	16.4	2.3	4.5	248.7	271.8
80–84	30.1	2.9	6.4	286.1	325.5
85+	46.6	2.5	7.9	279.6	336.6
Male	8.4	2.5	3.9	262.2	277.2
65–69	3.7	2.2	2.6	186.6	195.1
70–74	5.9	2.4	3.1	256.9	268.8
75–79	10.2	2.8	4.5	317.7	335.2
80–84	17.1	3.8	5.6	376.8	403.2
85+	27.9	4.7	8.6	389.4	430.5
Total	13.5	2.4	4.1	1212.4	257.0

[a]*ICD* discharge codes 800–904, 910–957.
[b]All nontrauma combined.

TABLE 4—Total Costs and Annual Per Capita Costs of Fall-Related Hospitalizations of Older Adults in Washington State

Costs	Female	Male	Total
Total costs, $ millions	37.98	15.36	53.35
Average cost per hospitalization, $	6559	7379	6776
Annual per capita cost, $, by age, y			
65–69	35	25	30
70–74	62	41	54
75–79	108	84	98
80–84	206	129	178
85+	310	200	279
Total	114	62	92

the frequency and cost of hospitalization for fall-related injuries in the older adult population of Washington State. These data confirm other research that implicates falls as the leading cause of injury in the aged.[2,5,13,14] The rates of such injuries in Washington were similar to national estimates reported by Rice et al.[13]

Fall-related injuries profoundly affect the activity and life-style of older adults. Forty-three percent of patients injured in falls were discharged to a nursing facility, a figure similar to that reported by Sattin et al.[15] Noninstitutionalized persons injured in falls were discharged to nursing care three times more often than were persons hospitalized for nontrauma causes.

Hospital costs for fall-related injuries in older adults is one of the more striking findings of this analysis. However, these hospitalized fall injuries account for only a fraction of the total cost.[13] Among an elderly population in Florida, only 42% of those who sought medical attention for a fall were hospitalized.[15] And the more than $53 million spent in hospital costs do not take into account the long-term sequelae of fall-related injuries, which include loss of independence, loss of confidence, susceptibility to future falls, and poor survivorship.[9,10] Clearly, the entire cost of falls among older adults in Washington State, whether measured in monetary or in human terms, is much greater than $92 per person per year (Table 4).

The Washington State Commission data have limitations typical of hospital discharge registries. Because the five discharge diagnosis codes are not prioritized, some falls may be secondary to comorbid conditions. Furthermore, E coding, though mandated by law, is not 100% complete; consequently, some of the injuries with unknown cause in fact may have been incurred by falls. Finally, no information is available regarding the course of the patient once discharged.

Although the etiology of falls is complex, fall injury research has identified several potentially modifiable risk factors[1–3,6,7,16] and fall prevention strategies.[5] Faulty vision, improper footwear, and common household hazards such as loose rugs are examples of previously identified risk factors. Currently unproven but potentially effective one-time interventions to reduce these risks deserve further attention. The cost-effectiveness of such interventions will increase over time as the risk and annual per capita cost of falls in older adults increase with age. These interventions may not prevent all falls, but

they could reduce the effect of falls on the older adult population. □

Acknowledgments

This study was supported by the Harborview Injury Prevention and Research Center, grant CCR 49-002570, Centers for Disease Control.

References

1. Hindmarah JJ, Estes EH. Falls in older persons. *Arch Intern Med.* 1989;149:2217–2222.
2. Perry BC. Falls among the elderly: a review of the methods and conclusions of epidemiologic studies. *J Am Geriatr Soc.* 1982;30:367–371.
3. Fisher ES, Baron JA, Malenka DJ, et al. Hip fracture incidence and mortality in New England. *Epidemiol.* 1991;2:116–122.
4. Grisso JA, Kelsey JL, Strom BL, et al. Risk factors for falls as a cause of hip fracture in women. *N Engl J Med.* 1991;324:1326–1331.
5. Tinetti ME, Speechley M. Prevention of falls among the elderly. *N Engl J Med.* 1989;320:1055–1059.
6. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. *N Engl J Med.* 1988; 319:1701–1707.
7. Ray WA, Griffin MR, Downey W. Benzodiazepines of long and short elimination half-life and the risk of hip fracture. *JAMA.* 1989;262:3303–3307.
8. Waller JA. Injury in aged: clinical and epidemiological implications. *N Y State J Med.* 1974;74(November):2200–2208.
9. Wild D, Nayak USL, Isaacs B. Prognosis of falls in old people at home. *J Epidemiol Community Health.* 1981;35:200–204.
10. Mossey JM, Mutran E, Knott K, Craik R. Determinants of recovery 12 months after hip fracture: the importance of psychosocial factors. *Am J Public Health.* 1989;79: 279–286.
11. Magaziner J, Simonsick EM, Kashner TM, Hebel JR, Kenzora JE. Survival experience of aged hip fracture patients. *Am J Public Health.* 1989;79:274–278.
12. From the WONDER computer network. Atlanta, Ga: Centers for Disease Control; 1991.
13. Rice DP, Mackenzie EJ, and Associates. Cost of injury in the United States: a report to Congress, 1989. San Francisco: Institute for Health and Ageing, University of California, San Francisco; Baltimore: Injury Prevention Center, School of Hygiene and Public Health, The Johns Hopkins University; 1989.
14. Centers for Disease Control. MMWR Surveillance of major causes of hospitalization among the elderly, United States, 1988. *MMWR CDC Surveill Summ.* April 1991.
15. Sattin RW, Lambert Huber DA, DeVito CA, et al. The incidence of fall injury events among the elderly in a defined population. *Am J Epidemiol.* 1990;131:1028–1037.
16. Nickens H. Intrinsic factors in falling among the elderly. *Arch Intern Med.* 1985; 145:1089–1093.

Carcinogens in Tobacco Smoke: Benzo[a]pyrene from Canadian Cigarettes and Cigarette Tobacco

Murray J. Kaiserman, PhD, and William S. Rickert, PhD

ABSTRACT

We evaluated the benzo[a]pyrene (BaP) content in the smoke from 35 brands of Canadian cigarettes and 5 brands of Canadian tobaccos for roll-your-own cigarettes. For the cigarettes, mean values of BaP ranged from 3.36 ng to 28.39 ng per cigarette, roughly in proportion with declared tar values. The relationship between declared tar and yields of BaP, however, does not allow accurate prediction of one from the other. For the tobaccos, mean BaP values ranged from 22.92 ng to 26.27 ng (average, 24.7 ng) per cigarette. The implications of these findings are discussed with respect to overall exposure. (*Am J Public Health.* 1992;82:1023–1026)

Introduction

The particulate fraction of tobacco smoke ("tar") is composed of a complex mixture of constituents,[1,2] some of which are regulated under environmental legislation.[3–6] One constituent is benzo[a]pyrene (BaP), a polynuclear aromatic hydrocarbon formed during the incomplete combustion of organic matter such as gasoline, garbage, and plants. BaP has been identified by the International Agency for Research on Cancer as an animal carcinogen and a probable human carcinogen (Class 2A)[7] with inhalation, oral ingestion, and dermal absorption as the important routes of entry.

Human data are unavailable, but oral ingestion of about 7 to 9 mg of BaP per kilogram has produced cancers in laboratory animals.[8] With respect to tobacco smoke, BaP has been detected in concentrations ranging from 20 to 40 ng per cigarette in mainstream cigarette smoke[1] to 40 to 79 ng per cigarette in sidestream smoke[1] and 96 to 292 ng per cigar in mainstream cigar smoke.[9]

On January 1, 1989, the Tobacco Products Control Act came into force in Canada. One of the purposes of this legislation is to "enhance public awareness of the hazards of tobacco use by ensuring the effective communication of pertinent information to consumers of tobacco products."[10] To achieve this goal, the Health Protection Branch of Health and Welfare Canada has undertaken a series of studies to develop and validate analytical test methods for toxic constituents.

In this paper we report on the BaP delivery of 35 brands of commercially available Canadian cigarettes and 5 brands

Murray J. Kaiserman is with the Tobacco Products Section, Environmental Health Directorate, Health and Welfare Canada, Ottawa, Ontario. William S. Rickert is president of Labstat Inc, Kitchener, Ontario.

Requests for reprints should be sent to Murray J. Kaiserman, PhD, Tobacco Products Section, Health and Welfare Canada, Ottawa, Ontario, Canada K1A 0L2.

This paper was submitted to the Journal March 4, 1991, and accepted with revisions November 27, 1991.

TABLE 1—Mean Benzo[a]pyrene (BaP) and Declared Tar, Nicotine, and Carbon Monoxide (CO) Concentrations per Cigarette in 36 Brands of Canadian Cigarettes

Brand	Tar, mg	Nicotine, mg	CO, mg	BaP, ng
Accord Ultra Mild KS	3	0.3	3	5.63
Black Cat Light RS FT	14	1.2	14	23.36
Black Cat RS FT	16	1.3	16	21.87
B&H 100's Light FT	10	1	10	16.86
Cameo Menthol KS FT	18	1.4	19	23.03
Craven A KS FT	13	1.2	13	22.53
Craven A Ultra Light KS FT	1	0.1	1	3.36
Du Maurier Extra Light RS FT	10	0.9	11	17.08
Du Maurier KS FT	16	1.3	17	22.92
Du Maurier Light KS FT	13	1.1	12	21.17
Du Maurier RS FT	15	1.2	16	23.06
Du Maurier Special KS FT	12	1	13	19.29
Export A Extra Light RS FT	8	0.9	9	17.66
Export A Light RS	12	1.3	12	17.8[illegible]
Export A Medium RS FT	13	1.2	14	19.04
Export A RS FT	15	1.3	16	25.86
Gitanes FT RS	12.9	0.89	16	12.40
Mark Ten FT KS	16	1.2	16	28.39
Mark Ten FT RS	17	1.3	17	24.02
Matinee Extra Mild KS FT	4	0.4	5	6.85
Matinee KS FT	13	1.1	12	13.16
Matinee Slims Extra Mild KS FT	3	0.4	3	8.21
Medallion KS FT Ultra Mild	1	0.2	2	4.60
Number 7 Light KS FT	13	1.1	13	20.19
Number 7 RS FT	16	1.3	16	19.74
Players Extra Light RS FT	11	1	11	15.65
Players Light KS FT	16	1.3	17	18.93
Players Light RS FT	16	1.3	15	15.49
Players RS FT	19	1.4	19	27.34
Rothmans Extra Light KS FT	10	1.1	9	10.49
Rothmans KS FT	16	1.2	16	17.69
Sportsmen FT KS	16	1.2	16	25.37
Vantage KS FT	9	1	11	17.30
Vantage Light KS FT	4	0.5	5	7.59
Viscount 1 Ultra Mild KS FT	0.7	0.09	0.7	3.39

Note. KS = king size; RS = regular size; FT = filter tip.

of commercially available Canadian fine-cut tobaccos for making roll-your-own cigarettes.

Methods

Smoking Procedures and Sample Collection

The 35 brands of cigarettes and 5 brands of fine-cut tobaccos were collected at the retail level in March and April of 1989 in Montreal and Vancouver. Brands were chosen to include some of the most popular brands in the market, including Players Light Regular Size (12.9% market share), Export A Regular Size (5.7% market share), and Rothmans King Size (3.7% market share).[11] The preparation of roll-your-own cigarettes and the smoking conditions for all cigarettes are described elsewhere.[12]

BaP was collected and analyzed according to the method described by Tomkins et al.[13] In this procedure, ^{14}C-labeled BaP is added to each sample and values, as determined by high-pressure liquid chromatography (HPLC), are corrected for ^{14}C recovery.

Recovery Study

Five Canadian control cigarettes (Monitor 4) were smoked per port on 10 randomly selected ports of a 20-port smoking machine and the mainstream particulate fraction was collected on a Cambridge filter pad, providing a total of 10 samples.

During cleanup, 100 µl of tracer solution was added to six of these samples at six different steps before HPLC isolation and 25 µl of tracer solution was added to two of the samples at two different steps after isolation and before HPLC quantitation. To determine 100% recovery, two of the samples were left unspiked (no tracer solution added) until just prior to injection onto the HPLC analytical column.

Validation Study

Five Kentucky Reference 2R1 cigarettes were smoked per port on 10 randomly selected ports of a 20-port smoking machine and the mainstream tar was collected on a Cambridge filter pad, resulting in another 10 samples. All samples were spiked with 100 µl of the tracer solution. Five of the pads were further spiked with 107.6 ng of unlabeled BaP. The workup then proceeded as described by Tomkins et al.[13]

Results and Discussion

Recovery and Validation Studies

The recovery and validation studies provide a measure of the precision, accuracy, and efficiency of the methods used. Observed recovery rates of BaP were 100% ± 8%. During the course of the experiment, the average recovery rate was 89% for the control cigarettes and 91% for fine-cut tobacco cigarettes; these rates were in good agreement with published results of 85% BaP recovery.[13] The results of the validation study provided a value of 31.9 ± 3.7 ng per cigarette for the unspiked 2R1 cigarette; this value was in good agreement with published values of 32 ng per cigarette[9] and 28.4 ± 3.4 ng per cigarette.[13]

Cigarette and Fine-Cut Tobacco Studies

Table 1 shows declared pack values for tar, nicotine, and carbon monoxide, and values for BaP as determined in this study. The results for BaP are averages of five observations per brand (25 cigarettes). (Although the current practice is to report yields on a per-cigarette basis, it should be pointed out that five cigarettes are required for each analytical value and the per-cigarette yield is obtained by division.)

Mean values for BaP range from 3.36 ng per cigarette to 28.39 ng per cigarette and are linearly related to declared tar values (correlation coefficient of 0.89 for all samples). Because the reported values reflect smoke concentrations under standard conditions, the smoker of low-yield cigarettes is probably exposed to higher amounts than those reported in Table 1.

Figure 1 shows the results of a regression analysis of mean BaP on declared tar. Although the simple linear correlation coefficient is fairly high at .89, the points are widely dispersed about the regression line, resulting in broad 95% confidence limits for the regression line. This means, for example, that at 16 mg declared tar, the BaP level is predicted to lie between 15 and 28 ng per cigarette. Thus, although

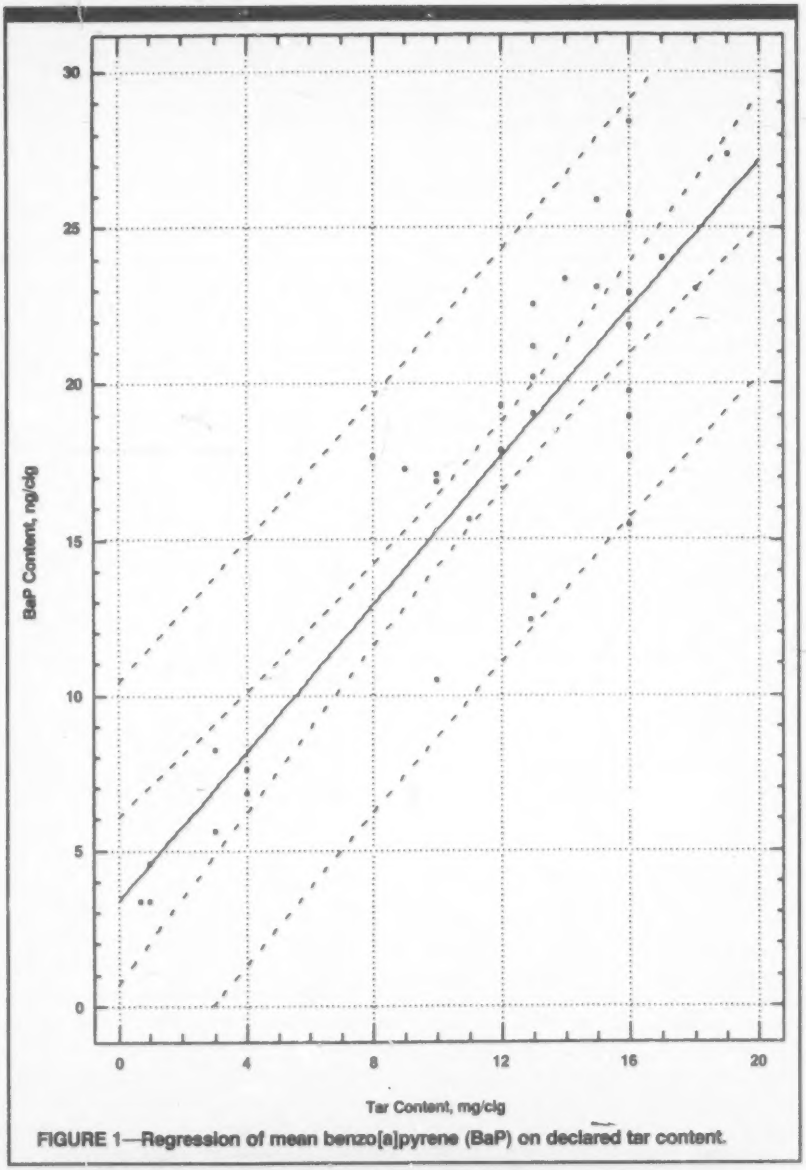


FIGURE 1—Regression of mean benzo[a]pyrene (BaP) on declared tar content.

TABLE 2—Mean Benzo[a]pyrene (BaP) and Declared Tar, Nicotine, and Carbon Monoxide (CO) Concentrations per Cigarette for Cigarettes Prepared from Five Brands of Roll-Your-Own Tobacco

Brand	Tar, mg	Nicotine, mg	CO, mg	BaP, ng
Players	23	1.8	19	25.02
Export	17	1.5	17	25.93
Craven A	19	1.5	19	23.34
Rothmans	19	1.5	19	26.27
Number 7	19	1.5	19	22.92

BaP is linearly related to declared tar, declared tar is not a good predictor of BaP.

Table 2 shows declared tar, nicotine, and carbon monoxide values and mean analytical BaP values for the five brands of fine-cut tobaccos used in preparation of roll-your-own cigarettes. Once again, BaP yields are averages of five observations per brand. These values range from 22.92 to 26.27 ng per cigarette; the average value was 24.7 ng per cigarette.

The Occupational Safety and Health Administration (OSHA) in the United States has set advisory levels for BaP of 0.2 $\mu g/m^3$ (8-hour time-weighted average permissible exposure limit for occupational exposure).[5] Assuming average breathing rates of 0.63 m^3/hour (women) to 1.08 m^3/hour (men) for alternating periods of light work and rest,[14] the permitted maximum amount of inhaled BaP in an occupational environment would be from 1.01 μg in 8 hours to 1.73 μg in 8 hours. A pack-a-day smoker who consumed 20 cigarettes in 8 hours could be expected to inhale from 0.067 μg to 0.568 μg from this source. Although the actual amount would vary depending upon individual smoking behavior, it is obvious that, for smokers, cigarette smoke is a significant source of exposure to BaP, contributing as much as one third the daily permitted exposure recommended by OSHA.

To clarify the risk further, the state of California mandates a "no significant risk" level for BaP of 0.06 μg per day.[6] A pack-a-day smoker (20 cigarettes) will inhale between one and ten times this level. For smokers, current advisory levels should be modified to take into account the contribution from tobacco smoke.

Although BaP is just one of 60 to 70 known toxic constituents of tobacco smoke,[1,2,7] many smokers continue to believe that tar is the most significant toxic constituent of tobacco smoke and that cigarettes with reduced tar are safer.[1,15] One reason for this misconception may be lack of knowledge of the components of tobacco smoke in a familiar context. □

References

1. US Department of Health and Human Services. *Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General.* US Dept of Health and Human Services, Public Health Service, Centers for Disease Control, Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. US Dept of Health and Human Services publication CDC 89-8411.
2. Dube MF, Green CR. Methods of collection of smoke for analytical purposes. *Recent Advances in Tobacco Science: Formation, Analysis and Composition of Tobacco Smoke.* 1982;8:42–102.
3. Canadian Environmental Protection Act. RSC C16 (4th suppl); 1988.
4. US Environmental Protection Agency. *Ambient Water Quality Criteria for Polynuclear Aromatic Hydrocarbons.* Washington, DC Office of Water Regulations and Standards; 1980. EPA document 440/5-80-069. NTIS no. PB81-117809.
5. US Occupational Safety and Health Administration. *Permissible Exposure Limits.* CFR 29:1910.1002 and 29:1910.1029; 1985.
6. 22 Cal Code of Regulations § 12711.
7. International Agency for Research on Can-

cer (IARC). *Monographs on the Evaluation of Carcinogenic Risks to Humans.* Lyon, France: IARC; 1987:58. Pub. Suppl 7.
8. *Toxicological Profile for Benzo[a]pyrene.* Atlanta, Ga: US Dept of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry; 1990. ATSDR/TP-85/05.
9. Appel BR, Guirguis G, Kim I-S, et al. Benzene, benzo(a)pyrene, and lead in smoke from tobacco products other than cigarettes. *Am J Public Health.* 1990;80:560–564.
10. The Tobacco Products Control Act. RSC C20; 1988.
11. Maxwell JC. How the brands ranked. *World Tobacco.* September 1990:121–124.
12. Kaiserman MJ, Rickert WS. Handmade cigarettes: it's the tube that counts. *Am J Public Health.* 1992;82:107–109.
13. Tomkins BA, Jenkins RA, Griest WH, Reagan RR, Holliday SK. Liquid chromatographic determination of benzo[a]pyrene in total particulate matter of cigarette smoke. *J Assoc Off Anal Chem.* 1985;68:935–940.
14. Rickert WS (Labstat Incorporated). Some considerations when estimating exposure to environmental tobacco smoke (ETS) with particular reference to the home environment. *Can J Public Health.* 1988;79: S33–S39.
15. Project Viking. "WAVE 2." Entered as evidence (AG-021D) in *RJR-MacDonald Inc v Attorney General of Canada* (Superior Court of Canada, District of Montreal). Court docket no. 500-05-009755-883; 1990.

Discrepancies between Self-reported Smoking and Carboxyhemoglobin: An Analysis of the Second National Health and Nutrition Survey

Lisa M. Klesges, MS, Robert C. Klesges, PhD, and Jeffrey A. Cigrang, MS

ABSTRACT

Environmental, self-report, and demographic factors mediated the relationship between self-reported cigarette smoking and carboxyhemoglobin among 2114 smokers and 3918 nonsmokers. Self-reported nonsmokers with carboxyhemoglobin levels between 2% and 3% were more likely to be self-reported ex-smokers, to live in a larger community, and to be younger, less educated, and male than were self-reported nonsmokers with carboxyhemoglobin levels of less than 2%. Self-reported nonsmokers with strong evidence of cigarette consumption (carboxyhemoglobin level > 3%) were more likely to be self-reported ex-smokers, younger, less educated, and non-White than were nonsmokers with carboxyhemoglobin levels of less than 2%. (*Am J Public Health.* 1992;82:1026–1029)

Introduction

Because of uncertainties inherent in measurement of smoking behavior using self-report,[1] biochemical measures are frequently used to validate reported smoking exposure. Although cotinine is often the preferred biochemical measure of smoking exposure,[2] blood carbon monoxide concentration is useful in many circumstances. It is less expensive to measure and can discriminate smokeless tobacco products from cigarette, pipe, and cigar smoking. It can also verify smoking cessation in treatment programs using nicotine replacement. Carbon monoxide is particularly useful in establishing health risks of smoking because of its strong relationship with cardiovascular disease.[3]

The current study sought to investigate factors associated with discrepancies between self-reported smoking status and carboxyhemoglobin levels. By identifying circumstances under which carboxyhemoglobin levels and self-reported smoking status do not agree, investigators can make more informed judgments about the validity of these measures of smoking and the associations between smoking and morbidity when carboxyhemoglobin is used.

Methods

The study sample consisted of 6032 participants from the Second National Health and Nutrition Examination Survey (NHANES II), aged 18 to 74 years, who completed a carboxyhemoglobin assessment. Detailed description of the survey methodology for NHANES II is available.[4] Self-reported smoking status was ascertained by asking participants whether they currently smoked cigarettes, how many they smoked on average, and whether they smoked cigars or pipes. Subjects reporting no current use of tobacco products were considered nonsmokers, while those smoking pipes or cigars were excluded. Carboxyhemoglobin concentrations, determined by spectrographic assay,[5] are expressed as percentages of total hemoglobin combined with carbon monoxide. Statistical analyses were performed using the Statistical Analysis System.[6] Collinearity diagnostics, as discussed by Hosmer and Lemeshow,[7] did not reveal unreliable estimates.

Lisa M. Klesges is with the Department of Biostatistics and Epidemiology, University of Tennessee, Memphis. Robert C. Klesges and Jeffrey A. Cigrang are with the Center for Applied Psychological Research, Department of Psychology, Memphis State University, Memphis, Tenn.

Requests for reprints should be sent to Lisa M. Klesges, MS, Department of Biostatistics and Epidemiology, University of Tennessee, Memphis, TN 38163.

This paper was submitted to the Journal February 4, 1991, and accepted with revisions December 31, 1991.

Results

Table 1 provides a demographic description of the study sample. Although the distributions overlapped, the average carboxyhemoglobin level of smokers was higher than that of nonsmokers. Based on its discriminant ability in previous investigations,[8,9] a carboxyhemoglobin level of 2% or more was used to identify smokers. Table 2 shows the relationship between self-reported smoking and elevated carboxyhemoglobin levels. Using carboxyhemoglobin as the standard, sensitivity and specificity were 84.8% and 95.8%, respectively, while the positive predictive value of carboxyhemoglobin was 91.5%. Two thirds of discrepancies involved subjects who reported that they smoked but had carboxyhemoglobin levels of less than 2%.

Subjects who reported smoking but had carboxyhemoglobin levels of less than 2% were categorized on cigarette rate (i.e., <10 cigarettes per day or 10+ cigarettes per day). Logistic regression equations compared each of these subgroups with self-reported smokers having carboxyhemoglobin levels of 2% or more. The first analysis compared the low cigarette consumption group with the accordant carboxyhemoglobin group. The odds of misclassification (see Table 3) were higher for Whites and those with lower cigarette consumption. Subjects who had carboxyhemoglobin levels of less than 2% and smoked 10+ cigarettes per day were also compared with smokers with accordant carboxyhemoglobin levels (≥2%). The odds of having discrepant vs accordant carboxyhemoglobin levels were increased for Whites, men, and those with lower cigarette consumption and higher body mass.

Among self-reported nonsmokers, logistic regression analyses were performed to identify factors that discriminated discrepant vs accordant subjects. Because marginally elevated carboxyhemoglobin levels may result from passive exposure and large elevations are more likely due to cigarette exposure, two logistic regression analyses were conducted. In the first, predictors of accordance (i.e., carboxyhemoglobin level < 2%) vs marginal discrepancies (i.e., carboxyhemoglobin level between 2% and 3%) were evaluated. In the second analysis, predictors of accordance vs severe discrepancies (i.e., carboxyhemoglobin level > 3%) were assessed.

Predictors of marginal discrepancies are presented in Table 4. Male subjects were more likely to be misclassified as smokers (61% more at risk) than were female subjects. Also, misclassification decreased with advancing age and education. The strongest predictor was size of the community of residence. The odds of having marginally discrepant carboxyhemoglobin levels increased threefold from rural to large urban (>3 million) areas. Finally, self-reported ex-smokers were 55% more likely to have discrepant carboxyhemoglobin levels than were never smokers.

The second logistic regression compared self-reported nonsmokers with carboxyhemoglobin levels of more than 3% and those with carboxyhemoglobin levels of less than 2%. Those with strong evidence of cigarette consumption were 75% more likely to be reported ex-smokers and twice as likely to be non-White compared with accordant nonsmokers. As age and education increased, the odds of having discrepant vs accordant carboxyhemoglobin levels decreased.

TABLE 1—Demographic Characteristics of the Subject Sample

	Self-reported Nonsmokers (n = 3918)	Self-reported Smokers (n = 2114)
Sex, No. (%)		
Male	1583 (40)	1123 (53)
Female	2335 (60)	991 (47)
Race, No. (%)		
Black	402 (10)	298 (14)
White	3444 (88)	1778 (84)
Other	72 (2)	38 (2)
Education, No. (%)		
High school or less	2204 (75)	1416 (67)
Some college	1714 (25)	698 (33)
Age, y		
Mean (SD)	48.3 (18.43)	42.4 (16.61)
Range	18–74	18–74
Carboxyhemoglobin level		
Mean (SD)	0.90 (0.96)	4.55 (2.35)
Range	0.00–11.53	0.05–12.92
Body mass index, kg/m^2		
Mean (SD)	25.79 (4.99)	24.62 (4.78)
Range	14.47–55.43	12.39–59.98
Size of community, No. (%)		
Rural	1411 (36)	720 (34)
Urban: 2500–999 999	1531 (39)	834 (39)
Urban: 1 million +	976 (25)	560 (26)

TABLE 2—Agreement of Self-reported Smoking Status and Carboxyhemoglobin Level

	Carboxyhemoglobin Level, No. (%)	
	< 2%	≥ 2%
Self-reported nonsmokers	3752 (95.8)	166 (4.2)
Self-reported smokers	321 (15.2)	1793 (84.8)

Discussion

The agreement between self-reported smoking and carboxyhemoglobin level of 91.9% is consistent with previous studies not using bogus pipeline or pipeline procedures to increase the validity of self-reports.[9,10] Discrepancies between self-reports and carboxyhemoglobin levels were not random but were predicted by demographic and smoking-related variables. The conundrum of these discrepancies is to differentiate falsification of self-reported smoking from failure of carboxyhemoglobin measures.

Because of this uncertainty, the role of biochemical verification in general surveys of smoking behavior remains controversial. Some data suggest increasing falsification in general surveys,[11,12] while other studies conclude that self-reports are superior to biochemical measures.[13] Self-reports of smoking in cessation pro-

TABLE 3—Predictors of Discrepancies in Self-reported Smokers

	< 10 Cigarettes per Day vs Accordant Carboxyhemoglobin Level (≥ 2%)[a]				≥ 10 Cigarettes per Day vs Accordant Carboxyhemoglobin Level (≥ 2%)[b]			
	Odds Ratio	(95% CI)	χ^2	*P*	Odds Ratio	(95% CI)	χ^2	*P*
No. of cigarettes (per 5 cigarettes/day)	14.78	(9.53, 22.91)	145.02	<.001	1.17	(1.08, 1.27)	13.77	<.001
Body mass, kg/m^2	0.95	(0.91, 1.00)	3.24	.072	0.95	(0.92, 0.98)	11.03	<.001
Race (0 = White, 1 = non-White)	3.16	(1.64, 6.11)	11.75	<.001	2.80	(1.58, 4.88)	12.71	<.001
Sex (0 = male, 1 = female)	0.83	(0.49, 1.41)	0.47	.491	0.68	(0.49, 0.93)	5.67	.017
Age (per 15 y)	0.96	(0.76, 1.22)	0.12	.729	0.94	(0.81, 1.09)	0.60	.437
Education (per 4 y)	0.79	(0.57, 1.10)	2.00	.157	0.90	(0.72, 1.12)	0.86	.354
Size of community (rural to 3 million +)	1.06	(0.95, 1.17)	1.06	.304	0.94	(0.89, 1.00)	3.82	.051

[a]n = 1912 (136 discrepancies).
[b]n = 1950 (176 discrepancies).

TABLE 4—Predictors of Discrepancies in Self-reported Nonsmokers

	Accordant (< 2%) vs Marginally Elevated Carboxyhemoglobin Level (2%–3%)[a]				Accordant (< 2%) vs Highly Elevated Carboxyhemoglobin Level (> 3%)[b]			
	Odds Ratio	(95% CI)	χ^2	*P*	Odds Ratio	(95% CI)	χ^2	*P*
Ex-smoker (0 = yes, 1 = no)	1.55	(0.98, 2.46)	3.48	.062	1.74	(1.07, 2.85)	4.92	.026
Race (0 = White, 1 = non-White)	0.94	(0.50, 1.76)	0.04	.846	2.00	(1.10, 3.66)	5.13	.023
Sex (0 = male, 1 = female)	0.62	(0.39, 0.97)	4.38	.036	0.71	(0.44, 1.15)	1.96	.161
Age (per 15 y)	0.78	(0.64, 0.94)	6.60	.010	0.69	(0.56, 0.85)	12.60	.001
Education (per 4 y)	0.74	(0.57, 0.95)	5.74	.017	0.73	(0.56, 0.96)	5.00	.025
Size of community (rural to 3 million +)	3.19	(1.64, 6.18)	11.80	.001	0.87	(0.42, 1.80)	0.13	.715

[a]n = 3810 (87 discrepancies).
[b]n = 3800 (77 discrepancies).

grams are inaccurate[14,15] as a result of interest in pleasing experimenters, experimental demand, and social pressures against smoking. Although falsification is reduced by bogus pipeline procedures[16] (i.e., subjects believe their reports will be biochemically validated), actual biochemical verifications are often recommended. Since, as the current findings suggest, carboxyhemoglobin level failed to predict the smoking status of certain individuals, blanket recommendations for carboxyhemoglobin verification are premature until the sources of disagreements between self-reports and carboxyhemoglobin levels are better understood. At present, the need for verification should be based on the characteristics of the target sample or cohort and the circumstances surrounding the collection of data.

Disagreement was highest (15.2%) in subjects who reported that they smoked but had carboxyhemoglobin levels of less than 2%. Discrepancies were predicted by the dosing variables of cigarettes per day and body mass, as well as race and sex. Because unmeasured smoking topography may differ by race and sex, it is difficult to distinguish physiological differences in carbon monoxide absorption and elimination from carbon monoxide exposure variables. However, regardless of causative factors, using carboxyhemoglobin level as the sole criterion of smoking for these individuals would bias estimates of disease risk related to smoking and the efficacy of treatment programs. Self-reports for these individuals are preferred over carboxyhemoglobin levels alone because the failure of carboxyhemoglobin is greater than that of self-reports in establishing smoking.

Self-reported nonsmokers with carboxyhemoglobin levels greater than 3% are most likely misreporting smoking because large elevations in carboxyhemoglobin levels attributed to carbon monoxide emissions are uncommon.[17] Self-reported exsmokers were likely to misreport smoking status, perhaps because of the social pressures of participating in a health examination survey. They may intend to quit smoking or have relapsed recently and want to claim ex-smoking status. Likewise, younger, non-White, and less-educated subjects misreported at a higher rate than their respective counterparts. These individuals may be responsive to perceived demand by interviewers or reticent to report smoking habits. Inhalation of carbon monoxide from other organic substances or occupational sources may be responsible for some carboxyhemoglobin elevations. Carbon monoxide in marijuana smoke reportedly increases false-positives.[18]

Finally, that marginal elevations in carboxyhemoglobin levels among self-reported nonsmokers may be due to passive exposure to carbon monoxide is supported by the finding that size of community predicted these marginal elevations. Marginal elevations were also predicted by being non-White and poorer educated, which may, in part, be explained by carbon monoxide emission in the home environment and occupational exposures[17] or a higher incidence of hemolytic anemia[19] in these individuals. □

Acknowledgments

This work was partially supported by Grants HL-45057 and HL-39332 from the National

Heart, Lung, and Blood Institute (Bethesda, Md) and by a Centers of Excellence grant awarded to the Department of Psychology, Memphis State University, by the state of Tennessee.

The authors are grateful to Dr. Stephen Kritchevsky and Dr. Judith Schwartzbaum for helpful advice on this manuscript and to the Biomedical Information Transfer Center for supplying computer services at the University of Tennessee, Memphis.

References

1. Pechacek TF, Fox BH, Murray DM, Luepker RV. Review of techniques for measurement of smoking behavior. In: Matarazzo JD, Weiss SM, Herd JA, Miller NE, Weiss SM, eds. *Behavioral Health: A Handbook of Health Enhancement and Disease Prevention.* New York, NY: Wiley; 1984:729–754.
2. Benowitz N. Biochemical measures of tobacco smoke consumption. In: Grabowski J, Bell C, eds. *Measurement of Smoking Behavior.* Washington, DC: National Institute on Drug Abuse; 1983. NIDA Research Monograph 48.
3. Wald NJ, Howard S, Smith PG, Kjeldsen K. Association between atherosclerotic diseases and carboxyhemoglobin levels in tobacco smokers. *Br Med J.* 1973;31:761–765.
4. National Center for Health Statistics. Plan and operation of the Second National Health-Nutrition Examination Survey, 1976–1980. Washington, DC: US Government Printing Office; 1982. DHHS publication no. (PH) 581-1317. (Vital and health statistics; series 1; no. 15).
5. Small KA, Radford EP, Frazier JM, Rodkey FL, Collison HA. A rapid method for simultaneous measurement of carboxy- and methemoglobin in blood. *J Appl Physiol.* 1971;31:154–160.
6. SAS Institute Inc. *SAS User's Guide: Statistics. Version 5 Edition.* Cary, NC: SAS Institute Inc; 1985.
7. Hosmer DW, Lemeshow S. *Applied Logistic Regression.* New York, NY: Wiley; 1989:131–132.
8. Vesey CJ, Salojee Y, Cole PV, Russell MAH. Blood carboxyhemoglobin, plasma thiocyanate, and cigarette consumption: implications for epidemiological studies of smokers. *Br Med J.* 1982;284:1516–1518.
9. Wald NJ, Idle M, Boreham J, Bailey A. Carbon monoxide in breath in relation to smoking and carboxyhemoglobin levels. *Thorax.* 1981;36:366–369.
10. Cohen JD, Bartsch GE. A comparison between carboxyhemoglobin and serum thiocyanate determinations as indicators of cigarette smoking. *Am J Public Health.* 1980; 70:284–286.
11. Vogt TM, Selvin S, Widdowson G, Hulley S. Expired air carbon monoxide and serum thiocyanate as objective measures of cigarette exposure. *Am J Public Health.* 1977; 67:545–549.
12. Warner K. Possible increases in the underreporting of cigarette consumption. *J Am Stat Assoc.* 1978;73:314–318.
13. Petitti D, Friedman G, Kahn W. Accuracy of information on smoking habits provided on self-administered research questionnaires. *Am J Public Health.* 1981;71:308–311.
14. Brockway BS. Chemical validation of self-reported smoking rates. *Behav Res Ther.* 1978;9:685–686.
15. Sillett R, Wilson M, Malcolm R, Ball K. Deception among smokers. *Br Med J.* 1978;2:1185–1186.
16. Jones EE, Sigal H. The bogus pipeline: a new paradigm for measuring affect and attitude. *Psychol Bull.* 1971;76:349–364.
17. Radford EP, Drizd TA. Blood carbon monoxide levels in persons 3–74 years of age: United States, 1976–1980. Hyattsville, Md: National Center for Health Statistics; 1982. NCHS advance data.
18. Pechacek T, Murray D, Luepker R, Mittelmark M, Johnson C, Schultz J. Measurement of adolescent smoking behavior: rationale and methods. *J Behav Med.* 1984; 7:123–140.
19. Stewart RD. The effect of carbon monoxide on humans. *Annu Rev Pharmacol.* 1975;15:409–425.

Recall of AIDS Public Service Announcements and Their Impact on the Ranking of AIDS as a National Problem

Michael Siska, MS, Janine Jason, MD, Paul Murdoch, MBA, Wen Shan Yang, PhD, and Robert J. Donovan, PhD

ABSTRACT

The efficacy of two public service announcements from Phase V of the "America Responds to AIDS" (ARTA) campaign was assessed at two sites. Participants were randomly assigned to view a local news program, one with an ARTA public service announcement appearing six times and the other with no AIDS public service announcements. During telephone interviews with 907 participants 1 to 3 nights after viewing, 21% at Site A and 59% at Site B could correctly recall the ARTA public service announcements. Absolute mentions of AIDS as an important national issue increased. (*Am J Public Health.* 1992;82:1029–1032)

Introduction

The evaluation of finished public health media material is often recommended for public health campaigns.[1–3] In late May and early June of 1990, the National AIDS Information and Education Program of the Centers for Disease Control ran a field study to test the short-term efficacy of two television public service announcements regarding human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS) from Phase V of the "America Responds to AIDS" (ARTA) campaign. The two public service announcements were "Wonderful World," which shows children playing in a field and a little girl talking about wanting to grow up in a better world, and "Sofa," which shows a young couple on a sofa kissing while an an-

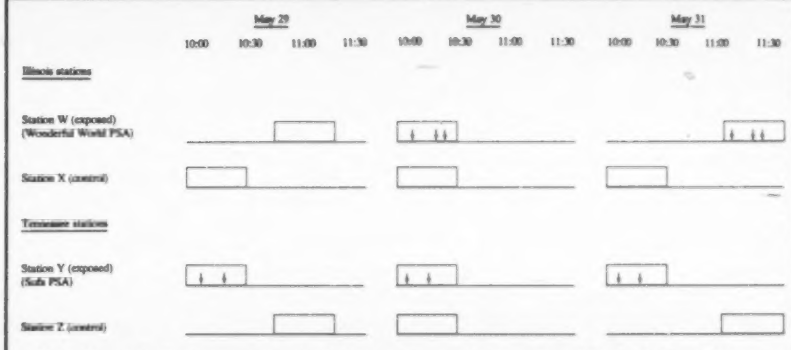
Michael Siska, Janine Jason, Wen Shan Yang, and Robert J. Donovan are with Applied Communications Research and Evaluation, National AIDS Information and Education Program, Centers for Disease Control, Department of Health and Human Services, Public Health Service, Atlanta, Ga. Paul Murdoch is with Ogilvy & Mather Advertising, Atlanta.

Requests for reprints should be sent to Michael Siska, MS, Mailstop E-25, NAIEP, CDC, 1600 Clifton Road, NE, Atlanta, GA 30333.

This paper was submitted to the Journal May 15, 1991, and accepted with revisions December 31, 1991.

TABLE 1—Number of Participants by Exposure Status and Study Site

	Springfield	Memphis	Total
Exposed	230	211	441
Unexposed	223	243	466
Total	453	454	907



Note. Boxes indicate approximate timing of newscasts. No other AIDS public service announcements were shown during any broadcasts on these stations. "↓" indicates the approximate time the ARTA public service announcement aired during that program.

FIGURE 1—Schedule of news and "America Responds to AIDS" (ARTA) public service announcement (PSA) airings.

TABLE 2—Demographic Characteristics of Study Participants by Exposure Category, Springfield

Demographic Category	Exposed, % (No.) (n = 230)	Control, % (No.) (n = 223)
Gender[a]		
Male	58 (134)	49 (109)
Female	42 (96)	51 (114)
Age		
18–24	10.5 (24)	10 (23)
25–34	28 (64)	28 (63)
35–44	34.5 (79)	37 (83)
45–54	27 (63)	24 (54)
Ethnic group		
White	93 (214)	95 (208)
Black	5 (12)	4 (8)
Other	2 (4)	1 (3)
Marital status		
Married	76 (175)	73 (160)
Never married	13 (29)	16 (36)
Divorced/separated/widowed	11 (26)	11 (25)
Education		
Less than high school	4 (10)	8 (18)
High school graduate	29 (66)	26 (58)
Some college	33 (76)	32 (71)
College graduate	18 (42)	20 (43)
Postgraduate education	16 (36)	14 (31)
Income		
< $15 000	10 (22)	11 (23)
$15 000–$35 000	42 (94)	40.5 (85)
$35 000–$65 000	41 (92)	37 (78)
$65 000+	8 (17)	11.5 (24)

[a]Significant at $P < .05$.

nouncer on the television in front of them talks about AIDS.

The study addressed four questions: (1) Did either public service announcement increase awareness of AIDS as an important national issue? (2) Did people remember the announcements? (3) Was one more effective than the other in accomplishing these two objectives? (4) Did either produce negative effects?

Methods

Each public service announcement was tested at a discrete site: "Wonderful World" was randomly assigned to the Springfield, Ill, "area of dominant influence," as defined by Arbitron, Inc,[4] and "Sofa" was randomly assigned to the Memphis, Tenn, area of dominant influence. A week prior to the study, 4181 eligible households were contacted via a random-digit dialing system. Of those contacted, 1909 (46%) declined to participate, 341 (8%) did not receive the channel they were asked to watch, and 1931 (46%) agreed to participate. An even gender distribution of participants 18 to 54 years old was randomly assigned to the exposed or unexposed (control) groups. Of the 1931 respondents agreeing to participate, 907 (47%) completed the follow-up questionnaire (Table 1).

At recruitment, participants were told only that the study was to examine important health and national issues facing the nation; AIDS was not mentioned. They were asked to watch one of two local late news programs and then asked to list "the most important problems facing our country today." The question on important health and national issues was also the first question asked during the follow-up interview to avoid any bias in responses. One station in each city aired no AIDS public service announcements (the control station). One station in each city was paid to air the announcement two times per night for 3 consecutive nights on its 10:00 PM news program, and each agreed to air no other AIDS public service announcements. The news programs were scheduled to air simultaneously, from 10:00 PM to 10:30 PM from May 29, 1990, through May 31, 1990, but the actual timing of the news programs and public service announcement airings was changed because of the National Basketball Association playoffs (Figure 1). After the scheduling problem was discovered, attempts were made to contact all of the study participants and ask for their continued cooperation. Follow-up telephone interviews were completed from June 1, 1990, through June 3, 1990.

TABLE 3—Demographic Characteristics of Study Participants by Exposure Category, Memphis

Demographic Category	Exposed, % (No.) (n = 211)	Control, % (No.) (n = 243)
Gender		
Male	45 (94)	47 (113)
Female	55 (117)	53 (130)
Age		
18–24	9.5 (20)	14 (34)
25–34	33 (70)	30 (73)
35–44	39 (82)	33 (81)
45–54	18.5 (39)	23 (55)
Ethnic group		
White	65 (137)	63 (151)
Black	34 (72)	36 (88)
Other	1 (2)	1 (3)
Marital status		
Married	63 (134)	65 (158)
Never married	25 (52)	21 (50)
Divorced/separated/widowed	12 (25)	14 (35)
Education		
Less than high school	9 (19)	11 (26)
High school graduate	23 (49)	27 (67)
Some college	39 (82)	35 (84)
College graduate	20 (43)	15 (37)
Postgraduate education	9 (18)	12 (29)
Income		
< $15 000	18 (36)	19 (45)
$15 000–$35 000	44 (90)	37 (87)
$35 000–$65 000	28 (57)	34 (80)
$65 000+	10 (21)	10 (23)

*Significant at $P < .05$.

TABLE 4—Proportion of Exposed Participants Recalling an AIDS Message, by Study Site

	Study Site, No. (%)	
	Springfield (n = 230)	Memphis (n = 211)
Unprompted	66 (24.4)	90 (42.7)
Prompted	47 (20.4)	63 (29.9)
Total	113 (44.8)	153 (72.6)

Note. Unprompted recall represents those respondents who replied by describing an AIDS message when asked if they had seen any public health messages. Prompted recall represents those who responded positively to the prompt "Did you see any commercial messages or public health messages that dealt with AIDS?" if they had not mentioned an AIDS message when asked if they had seen a public health message.

Results

The demographic characteristics of exposed vs control participants at both sites are presented in Tables 2 and 3.

The proportion of exposed participants recalling "an AIDS message" was significantly higher in the Tennessee site than in the Illinois site (Table 4; $P < .001$). In addition, 59% of those in Tennessee assigned to the exposure group could correctly recall specific elements of the "Sofa" public service announcement; 21% of those in Illinois assigned to the exposure group could correctly recall specific elements of "Wonderful World" ($P < .001$).

The proportion of participants mentioning AIDS as an important issue increased at both sites from recruitment to follow-up for the exposed groups, but not for the unexposed groups (Table 5). Mentions of AIDS increased 54.8% in Memphis (20.8% to 32.2%) and 72.9% in Springfield (9.6% to 16.6%). Categories of exposed participants vs unexposed participants were based strictly on recruitment assignment.

Three open-ended items were used to examine whether the public service announcements had unintended negative effects. In Illinois, 5 of 453 responses to an AIDS prevention question referred to "quarantine" of infected persons, 2 in the exposed group and 3 in the unexposed group. In Tennessee, 1 of 211 exposed respondents described the public service announcement as "offensive" and 1 described it as "too passionate to be shown on television." Since the general viewing population also could view the public service announcement, the station airing the announcement in Tennessee was questioned about complaints. The station received only 5 complaints, a number it classifies as negligible compared with a strong reaction, which it classifies as 25 or more calls and/or letters.

Discussion

A number of real-world problems and limitations were encountered in this study. First, unexpected scheduling changes occurred during the study. Second, all exposure information was self-reported. The validity of self-reported measures has been questioned,[5] but this method was the only practical one in this situation. Third, the overall response rate was low (47%), possibly because of the relatively large effort required of participants and the unforeseen scheduling changes. However, response rates to telephone surveys have been declining in recent years as a result of the proliferation of telephone survey research and telemarketing attempting to disguise itself as research.[6,7]

Day-after recall of an advertisement for a given product averages 20% (range = 0% to 70%).[8] Thus, the recall levels of "Wonderful World" (21%) and "Sofa" (59%) are comparable or superior to the recall levels obtained in commercial advertising.

The mention of AIDS as an important issue by exposed participants also increased at both sites. Baseline levels did differ between sites, and differed between groups in Springfield. Also, participants in Memphis may have been sensitized by an AIDS special that ran on a local news program 1 week prior to the study, which may account for baseline differences between sites. However, no stories on AIDS appeared on any of the news programs over the 3 nights. The public service announcements were the only AIDS-related materials aired. Furthermore, the increases were significant and occurred only among the exposed participants.

"Sofa" appeared more effective than "Wonderful World" in terms of viewer recall. Both appeared effective in prompting viewers to cite AIDS as an important issue facing the nation. However, direct

TABLE 5—Proportion and Change in Proportion Mentioning AIDS as an Issue, by Exposure Category

Category	Participants Mentioning AIDS, % Pre	Post	Change, %	P
Springfield				
Exposure	9.6	16.6	+ 7.00	.03
Control	17.5	17.5	...	NS
Memphis				
Exposure	20.8	32.2	+11.4	.01
Control	14.8	16.5	+ 1.6	NS

Note. At recruitment, and again during the follow-up interview, respondents were asked simply, "In your opinion, what are the most important problems facing our country today?" Up to the first five answers were recorded.

comparisons of the public service announcements are confounded by study site and airing schedule differences. It should also be noted that these results cannot be generalized to the entire viewing population.

Despite problems and limited exposure, this study shows that under field conditions viewers do recall ARTA public service announcements, and these announcements can affect viewers' beliefs concerning AIDS as a national problem. Furthermore, neither public service announcement produced negative effects.

The testing of finished public health media material is beneficial for practical decision making. For example, prior to this study, the Tennessee Department of Health had concerns that Tennessee residents would find the "Sofa" public service announcement offensive. However, the study results did not indicate this to be the case, and the announcement was later aired on a statewide basis.

Since ARTA campaign materials include print, radio, and television public service announcements and informational brochures and these materials are distributed extensively, and because this trial involved only six exposures over a short period of time, effects on attitudes and behaviors were not expected. However, future efficacy studies will be adapted to test for these effects. Upcoming efficacy studies will be run in several markets for 12 to 20 weeks following the release of the materials. They will examine various broadcast media combinations and media emphasis in relation to campaign objectives. □

Acknowledgments

This paper was presented at the 7th International AIDS Conference in Florence, Italy, June 1991.

The authors appreciate the assistance and advice of Chet Kelly and the Illinois Department of Public Health, Laurel Woods and the Tennessee Department of Public Health, and Mary Drew of Ogilvy & Mather Advertising.

References

1. Flay BR, Cook TD. Evaluation of mass media prevention campaigns. In: Rice R, Paisley W, eds. *Public Communication Campaigns.* Beverly Hills, Calif: Sage Publications; 1981:239–264.
2. DeJong W, Winsten JA. The use of mass media in substance abuse prevention. *Health Aff.* 1990;9:30–46.
3. Coyle SL, Boruch RF, Turner CF, eds. *Evaluating AIDS Prevention Programs.* Washington, DC: National Academy Press; 1991.
4. Taishoff S, ed. *The Broadcasting Yearbook: 1990.* New York, NY: Times Mirror; 1990.
5. Bauman KE, Dent CW. Influence of an objective measure on self-reports of behavior. *J Appl Psychol.* 1982;67:623–628.
6. Schleiter S. Trends in attitudes toward and participation in survey research. *Public Opinion Q.* 1986;50:17–26.
7. Rothenberg R. Surveys proliferate, but answers dwindle. *New York Times.* October 5, 1990:A1, A6.
8. Wells W, Burnett J, Moriarty S. *Advertising Principles and Practice.* Englewood Cliffs, NJ: Prentice-Hall; 1989.

NOTE

Some charts and graphs were color-coded in the original document. While it has not been possible to reproduce these items in color, their equivalent values on the black-and-white scale are often distinct enough to convey the same information as the original colors.

Congressional Information Service, Inc
4520 East-West Highway
Bethesda, Maryland 20814

NOTE

ONLY THE FOLLOWING PAGES OF THIS

PUBLICATION HAVE BEEN FILMED

Cov. 1, Pgs: 1065-1066,

1085-1157

A2623-1

American Journal of Public Health

August 1992, Vol. 82, No. 8 Established 1911

ED KIANKA 6-6389
CONGRESSIONAL INFORMATION SVC 028
4520 EAST-WEST HIGHWAY
BETHESDA MD 20814

APHA
American Public Health Association

Editorial

1079 Socioeconomic Differences in Mortality in Britain and the United States *G. Davey Smith and M. Egger*

Commentary

1082 National Mortality Rates: The Impact of Inequality? *R. G. Wilkinson*

Featuring Health Risks and Mortality

1085 Pregnancy-Related Mortality in New Jersey, 1975 to 1989 *K. J. Mertz, A. L. Parker, and G. J. Halpin*

1089 The Roseto Effect: A 50-Year Comparison of Mortality Rates *B. Egolf, J. Lasker, S. Wolf, and L. Potvin*

1093 The Quality of Life in the Year before Death *H. R. Lentzner, E. R. Pamuk, E. P. Rhodenhiser, R. Rothenberg, and E. Powell-Griner*

1099 HMO Membership, Treatment, and Mortality Risk among Prostatic Cancer Patients *H. P. Greenwald and C. J. Henke*

1105 Car Size or Car Mass: Which Has Greater Influence on Fatality Risk? *L. Evans and M. C. Frick*

1131 Socioeconomic Status and Childhood Mortality in North Carolina *M. D. Nelson, Jr*

1133 Does Equal Socioeconomic Status in Black and White Men Mean Equal Risk of Mortality? *J. E. Keil, S. E. Sutherland, R. G. Knapp, and H. A. Tyroler*

1136 Job Strain and Mortality in Elderly Men: Social Network, Support, and Influence as Buffers *A. Falk, B. S. Hanson, S.-O. Isacsson, and P.-O. Östergren*

1139 Mortality and the Attribution of Health Problems to Aging among Older Adults *W. Rakowski and T. Hickey*

1142 Use of Cardiovascular Disease Medications and Mortality in People with Older Onset Diabetes *B. E. K. Klein, S. E. Moss, and R. Klein*

1145 The Accuracy of the National Death Index When Personal Identifiers Other than Social Security Number Are Used *B. C. Williams, L. B. Demitrack, and B. E. Fries*

1147 Race and Sex Differences in Mortality following Fracture of the Hip *S. J. Jacobsen, J. Goldberg, T. P. Miles, J. A. Brody, W. Stiers, and A. A. Rimm*

Featuring Health Risks and Mortality

Annual Meeting Preliminary Program

(continued page 1065)

Contents

American Journal of Public Health

The Journal's one-page guidelines for authors, "What *AJPH* Authors Should Know," are printed in each issue. They can also be obtained from the Editorial Office.

APHA
American Public Health Association

August 1992, Vol. 82, No. 8 ISSN: 0090-0036 Established 1911

Articles

1113 The Development of Registries for Surveillance of Adult Lead Exposure, 1981 to 1992 *M. E. Baser*

1119 Lower Respiratory Illness in Infants and Low Socioeconomic Status *P. A. Margolis, R. A. Greenberg, L. L. Keyes, L. M. LaVange, R. S. Chapman, F. W. Denny, K. E. Bauman, and B. W. Boat*

1127 Imported Leprosy in the United States, 1978 through 1988: An Epidemic without Secondary Transmission *T. D. Mastro, S. C. Redd, and R. F. Breiman*

Public Health Briefs

1151 Seroprevalence of HIV-1 and Hepatitis B and C in Prostitutes in Albuquerque, New Mexico *S. R. Tabet, D. L. Palmer, W. H. Wiese, R. E. Voorhees, and D. R. Pathak*

1155 Decreased Access to Medical Care for Girls in Punjab, India: The Roles of Age, Religion, and Distance *B. E. Booth and M. Verma*

Health Law and Ethics

1158 HIV Prevention and the Two Faces of Partner Notification *R. Bayer and K. E. Toomey*

Notes from the Field

1165 Biotechnology Products in the Field: Bringing Regulation Closer to Home *R. S. Grossmann*

1166 Project Together: Serving Substance-Abusing Mothers and Their Children in Des Moines *E. Saunders*

New from NCHS

1168 Mortality among Minority Populations in the United States *L. A. Fingerhut and D. M Makuc*

Other Journal Departments

1066 APHA Masthead

1073 Advertisers' Index

1076 What *AJPH* Authors Should Know

1079 Journal Masthead

1088 Erratum

1171 Letters to the Editor

1178 Book Corner

1183 Job Opportunities

Preliminary Program of the APHA 1992 Annual Meeting

See page P-1 for contents

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.

APHA OFFICERS

Joyce C. Lashof, MD, *President*
C. William Keck, MD, MPH, *Immediate Past President*
Helen Rodriguez-Trias, MD, *President-Elect*
Arnold E. Greenberg, SM, *Vice-President (USA)*
Trevor Hancock, MB, MHSc, *Vice-President (Canada)*
Julio Teja Perez, MD, *Vice-President (Latin America)*
Caswell A. Evans, Jr, DDS, MPH, *Treasurer*
Frances M. Veverka, RS, MPH, *Chair, Executive Board*
George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*
William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD

Frances M. Veverka, RS, MPH (1992) *Chair*
Marilyn Aguirre-Molina, EdD, MS (1993)
Brian A. Dobrow, MPH (1992)
Caswell A. Evans, Jr, DDS, MPH, *Treasurer*
Kristine M. Gebbie, BSN, MN (1994) *Vice-Chair*
Alan R. Hinman, MD, MPH (1995)
Janet Heinrich, DrPH, MPH (1993)
Margaret H. Jordan, RN, MPH (1992)
C. William Keck, MD, MPH, *Immediate Past President*
Margo D. Kerrigan, MPH (1995)
Joyce C. Lashof, MD, *President*
Barry S. Levy, MD, MPH (1993)
Helen Rodriguez-Trias, MD, *President-Elect*
Allan G. Rosenfield, MD (1994)
Reed V. Tuckson, MD (1994)
Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO

William H. McBeath, MD, MPH, *Executive Director*
Joanne E. Lukomnik, MD, *Chair, Action Board*
George Pickett, MD, MPH, *Chair, Program Development Board*
George E. Hardy, MD, MPH, *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc, 1015 Fifteenth St, NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50 000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health,* 1015 Fifteenth St, NW, Washington, DC 20005.

Information on Manuscript Preparation is published monthly in the Journal. The Journal publishes reports of original research, research methods, program evaluations, analytic reviews, including health policy analysis, and reports for special departments. A paper is considered with the understanding that it or its essential substance neither has been published nor is under consideration elsewhere.

Subscription price, individual, is $100/year US, $140/year international; **institutional,** $160/year US, $200/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to *American Journal of Public Health,* 1015 Fifteenth St, NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St, NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St, NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS/COLLEAGUE, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc, Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Pregnancy-Related Mortality in New Jersey, 1975 to 1989

Kristen J. Mertz, MD, MPH, Artist L. Parker, MD, MPH, and George J. Halpin, MD, MPH

ABSTRACT

Objectives. National data are thought to underestimate pregnancy-related mortality in the United States. A multisource surveillance system for pregnancy-associated deaths in New Jersey offers an opportunity to identify the magnitude of and the trends in pregnancy-related mortality at the state level.

Methods. Data from all reported pregnancy-related deaths in the state from 1975 to 1989 were studied, and pregnancy mortality ratios were calculated.

Results. The New Jersey pregnancy mortality ratio decreased from the late 1970s to the early 1980s but began to rise in the late 1980s. The pregnancy mortality ratio for non-Whites was 3.6 times that for Whites for the 15-year period. The causes of pregnancy-related deaths changed over the 15-year period, with direct obstetrical causes playing a decreasing role. AIDS has become the major cause of pregnancy-related mortality in New Jersey. Finally, approximately 44% of the pregnancy-related deaths were considered to be preventable by the physician or patient or both.

Conclusions. New efforts must be made to combat the recent rise in pregnancy-related deaths, with special attention to preventing deaths among non-White women. (*Am J Public Health.* 1992;82:1085–1088)

Introduction

In studying deaths of pregnant and postpartum women, one statistical measure used by the Centers for Disease Control (CDC) is the pregnancy mortality ratio, the number of pregnancy-related deaths per 100 000 live births.[1] Pregnancy-related deaths include both direct deaths (which result from complications of pregnancy, labor, or delivery) and indirect deaths (which result from preexisting medical conditions that are exacerbated by pregnancy). Nonmaternal deaths, not included in the pregnancy mortality ratio, involve pregnant or postpartum women who die of causes unrelated to pregnancy (e.g., injuries, homicides) within the specified time frame. "Pregnancy-associated deaths" is an all-encompassing term that includes direct, indirect, and nonmaternal deaths.

Progress in preventing pregnancy-related deaths in recent decades has been impressive. National vital statistics data compiled by the National Center for Health Statistics (NCHS) show a decline in the pregnancy mortality ratio from 376 per 100 000 live births in 1940[2] to 7.8 per 100 000 in 1985.[3]

Despite this progress, pregnancy-related mortality remains a significant public health issue for several reasons. There is reliable evidence that national vital statistics data based solely on death certificates understate pregnancy mortality rates when compared with rates measured by data systems that supplement vital records with other forms of surveillance.[1,4,5] In addition, national rates are low only for Whites, with non-White rates being three times higher.[3,6] Furthermore, upon review, one third to one half of pregnancy-related deaths are considered preventable.[7,8,9] Finally, deaths from AIDS have begun to increase the pregnancy mortality ratio in some areas.[10]

Instituted in 1974, multisource surveillance and review of pregnancy-related deaths in New Jersey provide an opportunity to more accurately estimate the magnitude and trends in pregnancy-related mortality.[11] Death certificates, maternity service reports, physicians, other hospital personnel, medical examiners, and, more recently, computer searches of death tapes and hospital discharge tapes are all sources for the identification of pregnancy-associated deaths, which are reportable by law to the New Jersey Department of Health. The New Jersey Medical Society's Subcommittee on Maternal Mortality has collaborated with the New Jersey Department of Health since 1974 in an annual review of all cases.

Methods

The New Jersey Department of Health maintains confidential case files—which include death certificates, hospital records, prenatal charts, and autopsy reports—on all women who died while pregnant or within 42 days postpartum. We compiled a database that included variables on demographic information, pregnancy history, social history, delivery, and fetal outcome. Cause of death, classification (direct, indirect, or nonma-

The authors are with the Division of Family Health Services, New Jersey State Department of Health, Trenton, NJ.

Requests for reprints should be sent to George J. Halpin, MD, MPH, New Jersey State Department of Health, 363 West State St, Trenton, NJ 08625.

This paper was submitted to the Journal April 18, 1991, and accepted with revisions April 7, 1992.

TABLE 1—New Jersey Pregnancy Mortality Ratios, 1975 to 1989

	Whites Ratio (No.)	Non-Whites Ratio (No.)	Total Ratio[a] (No.)[b]
Five-year period of death			
1975–1979	15.3 (55)	62.1 (60)	24.7 (115)
1980–1984	12.7 (48)	29.8 (31)	16.1 (79)
1985–1989	11.9 (51)	52.4 (64)	20.3 (115)
Age group			
≤19	12.1 (11)	31.2 (24)	21.4 (35)
20–24	11.4 (32)	31.9 (32)	18.3 (64)
25–29	11.6 (49)	50.0 (40)	17.9 (89)
30–34	12.6 (35)	77.2 (36)	22.5 (71)
35–39	18.7 (15)	92.8 (15)	35.8 (30)
≥40	108.0 (12)	286.0 (8)	153.0 (20)

Note. The pregnancy mortality ratios were calculated from data collected on the whole population or entire subpopulations of New Jersey, not on a sample of the population, and thus confidence intervals are not necessary.
[a]Number of pregnancy-related deaths per 100 000 live births. The number of live births includes multiple births and omits miscarriages, fetal deaths, and stillbirths.
[b]Number of pregnancy-related deaths.

ternal), and preventability had been determined by consensus of the Maternal Mortality Subcommittee of the New Jersey Medical Society, and these variables were added to the database.

Pregnancy mortality ratios were calculated by year, race, age, county, insurance status, cause of death, and preventability. The denominator data (numbers of live births by year, race, age, and county) were provided by the New Jersey Department of Health's Center for Health Statistics.[12] Hispanic live births were grouped with Whites until 1989 and thus pregnancy mortality ratios could only be calculated for Whites and Hispanics combined (Whites) and for non-Hispanic Blacks, Asians, Native Americans, and all other races combined (non-Whites). The data on number of live births by insurance status for the years 1985 to 1988 were provided by the New Jersey Department of Health's Division of Health Planning and Resource Development.

Results

Of the 404 pregnancy-associated deaths reviewed by the New Jersey Medical Society, seven were not classified and 88 were classified as nonmaternal and will not be considered here. The remaining 309 pregnancy-related deaths give an overall pregnancy mortality ratio of 20.3 deaths per 100 000 live births in New Jersey for the years 1975 to 1989. The general trend in the ratio in New Jersey was downward until it began to increase after 1985, rising to a high of 33.2 in 1988. The 5-year pregnancy mortality ratio varied from 24.7 for 1975 to 1979, 16.1 for 1980 to 1984, and 20.3 for 1985 to 1989 (Table 1).

Demographic Differences

The total 15-year pregnancy mortality ratio for non-Whites was 48.0 per 100 000 live births, 3.6 times as high as the ratio for Whites (13.2 per 100 000). After a substantial decrease between the periods 1975 to 1979 and 1980 to 1985 (62.1 to 29.8), the pregnancy mortality ratio for non-Whites increased to 52.4 in 1985 to 1989, whereas the ratio for Whites continued to decline from 15.3 to 12.7 to 11.9 over the same time periods (Table 1). The non-White to White ratio of race-specific pregnancy mortality ratios changed from 4.1 for 1975 to 1979 to 2.3 for 1980 to 1984 and 4.4 for 1985 to 1989.

The pregnancy mortality ratios by age group show that the pregnancy mortality risk in New Jersey during 1975 to 1989 was far higher for women aged 40 or older (Table 1). Women aged 20 to 24 and 25 to 29 had the lowest pregnancy mortality ratios (18.3 and 17.9, respectively). Within each age group the ratio for non-Whites was higher than that for Whites.

The pregnancy mortality ratios varied considerably by county, with the highest ratios in urban counties with large non-White populations. Essex County had the highest ratio, with 38.3 pregnancy-related deaths per 100 000 live births, followed by Hudson, Mercer, and Union Counties at 31.9, 26.4, and 25.5, respectively. The county-specific pregnancy mortality ratios were correlated with county-specific non-White population percentages (Pearson correlation coefficient = .84). In 15 of the 21 counties, the pregnancy mortality ratio for non-Whites was higher than the ratio for Whites. The remaining six counties have small non-White populations and had pregnancy mortality ratios of zero for non-Whites.

Pregnancy mortality ratios by insurance coverage could only be calculated for 1985 to 1988 because denominator data were available for only these years. Of the 97 pregnancy-related deaths that occurred during these 4 years, insurance status was recorded for 74 (73%) of them. These 74 accounted for pregnancy mortality ratios of 10.8 for those women with private insurance, 34.3 for women covered by Medicaid, and 32.7 for self-pay patients. Insurance status for live births was not available by race, so pregnancy mortality ratios by both payer and race could not be calculated.

Cause of Death

The underlying cause of death was obtained from the death certificate and presumably reflected findings on autopsy, performed in about 71% of the cases. The leading causes of pregnancy-related mortality in New Jersey changed significantly over the course of the three time periods. The cause-specific pregnancy mortality ratios for hypertension in pregnancy, thromboembolism, hemorrhage, and stroke (the leading causes of death in 1975 to 1979) all decreased in the 1980s (Table 2). The pregnancy mortality ratio for AIDS increased from 0.0 before 1985 to 2.3 per 100 000 in 1985 to 1989, the highest cause-specific ratio of that period. Other indirect causes also played an increasing role in 1985 to 1989. From 1975 to 1989, the cause-specific pregnancy mortality ratios were higher for non-Whites than Whites for every major cause of death (Table 2).

Preventability

The New Jersey Medical Society's Subcommittee on Maternal Mortality assigned preventability status to all but 7% of the pregnancy-related deaths for 1975 to 1989. Of those categorized, 56% were considered to be nonpreventable, 34% were considered to be preventable by physician factors, 5% preventable by patient factors, and 5% preventable by both patient and physician factors. Thus, for all deaths for which preventability status was assigned, 44% were considered preventable by the physician or patient or both.

The pregnancy mortality ratio for preventable deaths with a component of physician responsibility decreased from

the first time period to the second but showed no further decrease in the third (ratios of 9.5, 6.3, and 6.4, respectively) (Table 3). However, for non-Whites the pregnancy mortality ratio for physician-preventable deaths increased in the second half of the 1980s, paralleling the increase in pregnancy-related deaths as a whole for that period. For the total 15-year period, the pregnancy mortality ratio for physician-preventable deaths was 4.8 times as high for non-Whites as it was for Whites (19.5 per 100 000 compared with 4.1 per 100 000).

TABLE 2—New Jersey Cause-Specific Pregnancy Mortality Ratios, 1975 to 1989

Cause of Death	All Races						Whites		Non-Whites	
	1975–1979 Ratio[a]	(No.)[b]	1980–1984 Ratio	(No.)	1985–1989 Ratio	(No.)	1975–1989 Ratio	(No.)	1975–1989 Ratio	(No.)
Thromboembolism	2.6	(13)	2.0	(10)	1.2	(7)	1.8	(21)	2.8	(9)
Amnionic fluid embolism	1.1	(5)	1.8	(9)	0.9	(5)	0.7	(8)	3.4	(11)
Hypertension-in-pregnancy	3.7	(17)	0.8	(4)	1.6	(9)	1.1	(13)	5.3	(17)
Ectopic pregnancy	1.5	(7)	0.6	(3)	1.8	(10)	0.6	(7)	4.0	(13)
Hemorrhage	2.4	(11)	1.8	(9)	1.6	(9)	0.7	(8)	6.5	(21)
Anesthesia complication	1.9	(9)	0.8	(4)	1.6	(9)	0.8	(9)	4.0	(13)
Abortion-related	0.4	(2)	0.4	(2)	0.7	(4)	0.1	(1)	2.2	(7)
Obstetrical infection	0.9	(4)	1.0	(5)	0.0	(0)	0.5	(6)	0.9	(3)
Cerebrovascular accident	2.2	(10)	1.6	(8)	0.4	(2)	1.3	(15)	1.5	(5)
AIDS-related	0.0	(0)	0.0	(0)	2.3	(13)	0.6	(7)	1.9	(6)
Cancer	0.6	(3)	0.0	(0)	0.4	(2)	0.3	(4)	0.3	(1)
Other direct	1.5	(7)	0.8	(4)	0.9	(5)	0.5	(6)	3.1	(10)
Other indirect	4.9	(23)	3.9	(19)	6.5	(37)	3.9	(46)	10.2	(33)
Unexplained	0.9	(4)	0.4	(2)	0.5	(3)	0.3	(3)	1.9	(6)
Total	24.7	(115)	16.1	(79)	20.3	(115)	13.2	(154)	48.0	(155)

[a]Number of pregnancy-related deaths per 100 000 live births.
[b]Number of pregnancy-related deaths.

TABLE 3—New Jersey Pregnancy Mortality Ratios For Physician-Preventable Deaths, By Race, 1975 to 1989

Time Period	Total Physician-Preventable Ratio[a]	(No.)[b]	Whites Physician-Preventable Ratio	(No.)	Non-Whites Physician-Preventable Ratio	(No.)
1975–1979	9.5	(44)	5.0	(18)	26.9	(26)
1980–1984	6.3	(31)	4.2	(16)	14.4	(15)
1985–1989	6.4	(36)	3.3	(14)	18.0	(22)
Total	7.3	(111)	4.1	(48)	19.5	(63)

[a]Number of pregnancy-related deaths per 100 000 live births.
[b]Number of pregnancy-related deaths.

Discussion

Several findings in this study are of particular interest. First, we found that pregnancy mortality ratios for New Jersey are higher than the national ratios described by the NCHS, the CDC, and the Maternal Mortality Collaborative.[3,5,6,13,14] However, a comparison of New Jersey and national data may be misleading because of different systems of surveillance and different definitions of pregnancy-related mortality. Surveillance with death certificates alone leads to underestimation of pregnancy-related mortality [4,5,11,15] New Jersey's multiple-source surveillance system allows the state to identify pregnancy-related deaths that might not be identified by the NCHS and the CDC in other jurisdictions.

On the other hand, New Jersey uses a 42-day postpartum period in its definition, whereas the Maternal Mortality Collaborative and the CDC use a one-year postpartum period. Of the deaths reported to the Maternal Mortality Collaborative, 11% of postpartum deaths took place after 42 days postpartum.[5] Thus New Jersey data may underestimate the number of pregnancy-related deaths when compared with the Maternal Mortality Collaborative and the CDC. A uniform methodology for identifying, defining, and classifying pregnancy-associated deaths will have to be developed for use throughout the country before meaningful comparisons can be made.

The second issue of concern is the exceptionally high pregnancy mortality ratio for non-Whites in New Jersey (48.0/100 000). This ratio far exceeds the US Public Health Service's health objective for 1990 of a maternal death rate of less than 5 per 100 000 for any ethnic group in the United States.[16] Of particular concern is the high rate of physician-preventable deaths for non-White women as compared with that for White women, which suggests a possible discrepancy in the quality of obstetrical care provided to Whites and non-Whites. Special efforts must be made to improve obstetrical care, access to prenatal care, and family planning services in order to reduce the number of pregnancy-related deaths among non-White women.

Also of concern are the high pregnancy mortality ratios for women on Medicaid or without any health insurance coverage. The pregnancy mortality ratios by payer calculated for 1985 to 1988 underestimate the true ratios for this period because of the exclusion of several pregnancy-related deaths for which insurance status was not known. In addition, we could not analyze the payer data by race. The interaction between race, health insurance status, and poverty should be studied in greater detail in an attempt to separate individual effects.

The data also show that indirect causes of pregnancy-related death, such as AIDS, are replacing direct obstetrical causes in New Jersey. This trend has implications for the practice of obstetrics and the design of regional perinatal systems. Perinatal providers must now involve other medical personnel, such as infectious disease specialists, in the management of pregnancy in order to reduce mortality from indirect causes such as AIDS. Toward this end, perinatal systems will have to include a wider base of consultants to care for pregnant women with underlying disease. In addition, obstetricians and other health providers, through preconception counseling, must emphasize the danger of pregnancy when a significant underlying disease exists and help provide access to contraceptive services. □

Acknowledgments

We are grateful to Dr. Zena Stein and Deborah Maine of the School of Public Health, Columbia

University, for their review and comments. We also thank Dr. James Thompson of the Maternal Mortality Subcommittee, New Jersey Medical Society, for his support and assistance with the data collection.

References

1. Ellerbrock TV, Atrash HK, Hogue CJR, Smith JC. Pregnancy mortality surveillance: a new initiative. *Contemp Ob/Gyn.* June 1988:23–24.
2. *Vital Statistics of the United States, 1940, Part II.* Washington, DC: Bureau of the Census; 1943.
3. Advance report of final mortality statistics, 1985. *Monthly Vital Stat Rep.* Hyattsville, Md: National Center for Health Statistics; 1986;36(suppl):41.
4. Smith JC, Hughes JM, Pekow PS, Rochat RW. An assessment of the incidence of maternal mortality in the United States. *Am J Public Health.* 1984;74:780–783.
5. Rochat RW, Koonin LM, Atrash HK, Jewett JF. Maternal mortality in the United States: report from the Maternal Mortality Collaborative. *Obstet Gynecol.* 1988;72: 91–97.
6. Kaunitz AM, Hughes JM, Grimes DA, Smith JC, Rochat RW, Katrissen ME. Causes of maternal mortality in the United States. *Obstet Gynecol.* 1985;65:605–612.
7. Gibbs CE, Locke WE: Maternal deaths in Texas, 1969 to 1973. *Am J Obstet Gynecol.* 1976;126:687–692.
8. Sachs BP, Brown DAJ, Driscoll SG, et al. Maternal mortality in Massachusetts. *N Engl J Med.* 1987;316:667–672.
9. Hansen GF, Noone TA. Maternal deaths in New Jersey. *NJ Med.* 1988;85:1020–1022.
10. Koonin LM, Ellerbrock TV, Atrash HK, et al. Pregnancy-associated deaths due to AIDS in the United States. *JAMA.* 1989; 261:1306–1309.
11. Ziskin LZ, Gregory M, Kreitzer M. Improved surveillance of maternal deaths. *Int J Gynaecol Obstet.* 1979;16:281–286.
12. *New Jersey Health Statistics.* Trenton, NJ: New Jersey Department of Health, Center for Health Statistics; 1975–1989.
13. Advance report of final mortality statistics, 1978. *Monthly Vital Stat Rep.* Hyattsville, Md: National Center for Health Statistics; 1980;29(suppl):37.
14. Advance report of final mortality statistics, 1980. *Monthly Vital Stat Rep.* Hyattsville, Md: National Center for Health Statistics; 1983;32(suppl):35.
15. Rubin G, McCarthy B, Shelton J, Rochat RW, Terry J. The risk of childbearing reevaluated. *Am J Public Health.* 1981;71: 712–716.
16. *Promoting Health/Preventing Disease: Objectives for the Nation.* Atlanta, Ga: US Public Health Service; 1980.

Erratum

In: Feldman JG, Minkoff HL, McCalla S, Salwen M. A cohort study of the impact of perinatal drug use on prematurity in an inner-city population. *Am J Public Health.* 1992;82:726–728.

In Table 1, the confidence interval for prenatal care should have read ".21–.63." In Table 2, the confidence interval for hypertension should have read "1.99–14.28," and the confidence interval for prenatal care, ".22–.89."

The Roseto Effect: A 50-Year Comparison of Mortality Rates

ABSTRACT

Objectives. Earlier studies found striking differences in mortality from myocardial infarction between Roseto, a homogeneous Italian-American community in Pennsylvania, and other nearby towns between 1955 and 1965. These differences disappeared as Roseto became more "Americanized" in the 1960s. The present study extended the comparison over a longer period of time to test the hypothesis that the findings from this period were not due to random fluctuations in small communities.

Methods. We examined death certificates for Roseto and Bangor from 1935 to 1985. Age-standardized death rates and mortality ratios were computed for each decade.

Results. Rosetans had a lower mortality rate from myocardial infarction over the course of the first 30 years, but it rose to the level of Bangor's following a period of erosion of traditionally cohesive family and community relationships. This mortality-rate increase involved mainly younger Rosetan men and elderly women.

Conclusions. The data confirmed the existence of consistent mortality differences between Roseto and Bangor during a time when there were many indicators of greater social solidarity and homogeneity in Roseto. (*Am J Public Health.* 1992; 82:1089–1092)

Brenda Egolf, MA, Judith Lasker, PhD, Stewart Wolf, MD, and Louise Potvin, PhD

Introduction

Previous studies of social correlates of health and longevity have focused mainly on ethnic, occupational, or demographic characteristics rather than on the social character and group dynamics of an identified community. An exception is the study of Roseto, Pennsylvania, an Italian-American town in eastern Pennsylvania that from 1955 to 1965 was found to have had a strikingly low mortality rate from myocardial infarction relative to Bangor, an immediately adjacent town, and three other nearby communities.[1,2] The usually accepted risk factors were at least as prevalent in Roseto as in the two control communities that were studied in detail.[3,4] At the time, Roseto, which had been settled by immigrants from a town in southern Italy in 1882, still displayed a high level of ethnic and social homogeneity, close family ties, and cohesive community relationships.

Beginning in 1962, Roseto became the object of detailed inquiries into its history and social traditions and the attitudes, behavior, health habits, and medical status of its residents. For comparison, the adjacent community of Bangor (which, throughout what became a 50-year period under study, was served by the same water supply, physicians, and hospital facilities as Roseto) and the nearby town of Nazareth were studied in similar fashion. Data from these studies and many years of participant observation yielded evidence that the discrepancy in the prevalence of and mortality from myocardial infarction might be attributed to important differences in culture and social cohesion among the three communities.[1–5] It was hypothesized that Roseto's stable structure, its emphasis on family cohesion, and the supportive nature of the community may have been protective against heart attacks and conducive to longevity. However, by the early 1960s there were indications of impending change in the community.[5] Interviews with many Rosetans between the ages of 25 and 35 and with the teenagers indicated that they were prepared to largely abandon their old community ways in favor of the more typically American behavior of neighboring communities.

In 1963, after the initial period of study, the investigators made a prediction that the loosening of family ties and community cohesion would be accompanied by loss of relative protection of Rosetans from death due to myocardial infarction.[4] By the late 1960s and early 1970s the predicted social change was evident, as was the predicted increase in incidence of myocardial infarction. These changes were reflected in greater prevalence of myocardial infarction among the living and in higher mortality rates.[2,13,14]

The earlier beliefs and behavior that expressed themselves in Roseto's family-centered social life, absence of ostentation even among the wealthy, nearly exclusive patronage of local business, and a predominance of intra-ethnic marriages gradually changed toward the more familiar behavior pattern of neighboring commu-

Brenda Egolf and Judith Lasker are with the Center for Social Research, Lehigh University, Bethlehem, Pa. Stewart Wolf is with the Totts Gap Medical Laboratory, Bangor, Pa. Louise Potvin is with the University of Montreal Medical School, Montreal, Quebec, Canada.

Requests for reprints should be sent to Brenda Egolf, MA, Center for Social Research, Lehigh University, 203 East Packer Ave, Bethlehem, PA 18105.

This paper was submitted to the Journal March 8, 1991, and accepted with revisions December 10, 1991.

TABLE 1—Age-Adjusted Mortality Rates (MR) per 1000 and Number of Deaths for Roseto and Bangor Residents over 35 Years of Age

	1935–1944		1945–1954		1955–1964		1965–1974		1975–1984	
	MR	No. of Deaths	MR	No. of Deaths	MR	No. of Deaths	MR	No. of Deaths	MR	No. of Deaths
Myocardial infarction										
Bangor men	39.7	49	64.9	94	75.2	118	82.6	137	76.3	124
Roseto men	23.0	6	43.0	13	49.6	19	94.6	36	78.5	38
Bangor women	23.5	30	27.4	48	38.5	78	33.6	88	34.5	101
Roseto women	12.5	3	19.1	6	25.0	10	36.1	18	36.6	27
Congestive heart failure										
Bangor men	53.9	67	37.4	56	28.6	47	34.7	60	16.9	27
Roseto men	43.9	11	31.3	10	25.3	10	29.5	11	8.6	5
Bangor women	50.3	64	31.3	56	29.7	65	23.9	64	15.1	51
Roseto women	52.9	12	42.3	13	23.5	10	26.5	13	13.2	11
Myocardial infarction and congestive heart failure										
Bangor men	93.5	116	102.3	150	103.8	165	117.2	197	93.2	151
Roseto men	66.8	17	74.3	23	74.9	29	124.1	47	87.0	43
Bangor women	73.8	94	58.6	104	68.2	143	57.5	152	49.6	162
Roseto women	65.4	15	61.4	19	48.5	20	62.6	31	49.8	38
All causes										
Bangor men	242.0	299	209.8	306	217.8	346	244.1	406	222.4	351
Roseto men	186.9	48	179.9	57	174.0	66	262.8	100	172.5	82
Bangor women	202.0	257	141.6	247	158.0	326	127.7	317	124.3	359
Roseto women	193.0	46	159.5	50	125.1	51	150.0	75	93.9	70

nities. Roseto was shifting from its initially highly homogeneous social order—made up of three-generation households with strong commitments to religion and to traditional values and practices—to a less cohesive, materialistic, more "Americanized" community in which three-generation households were uncommon and inter-ethnic marriages became the norm.[15]

The "Roseto Effect" has been widely cited as evidence for the positive effects of social cohesion and social support on longevity.[6–12] In order to reexamine the possibility of bias in the mortality data due to the small population size of Roseto (approximately 1600 during the time of the study, with the Bangor population numbering more than 5000), and to test the possibility that the relative differences in mortality rates from one decade to another were due to random fluctuations in the number of deaths in each town, we examined 3859 death certificates from Roseto and Bangor from 1935 to 1985.

Method

With the help of the Bureau of Vital Statistics in the Pennsylvania Department of Health, we were able to obtain and individually examine all death certificates of inhabitants of Roseto and Bangor who died between 1935 and 1985. Because pre-1960 certificates had to be retrieved manually, we were concerned about the completeness of our data. We therefore examined all obituary notices in the local newspaper as well as church death records for the entire time period of this study; we were thus able to identify and obtain the records of residents whose certificates had not been filed or retrieved appropriately. Obituaries and church records also made it possible to identify and obtain the death certificates for Roseto and Bangor residents who had died while they were traveling outside of Pennsylvania.

Diagnostic Criteria

The criteria for the diagnosis of death from myocardial infarction included all instances when myocardial infarction was listed on line one of the death certificate as "immediate cause of death" or when "sudden death" was listed in that space and myocardial infarction, ischemic heart disease, arteriosclerotic cardiovascular disease, coronary heart disease, or similar phrases were used in the "due to" lines.

Criteria for the diagnosis of congestive heart failure included the listing of congestive heart failure, chronic ischemic heart disease, chronic myocarditis, cardiomyopathy, and similar phrases as immediate cause of death, accompanying the listing of atherosclerosis or arteriosclerosis as a contributing cause.

Age Adjustment and Standardization

Mortality rates in the two towns were standardized for age, using as the standard population the sum in each age group in Roseto and Bangor in 1940. Rates were computed for deaths from myocardial infarction, congestive heart failure, and all causes. The ratio of Roseto to Bangor rates was then calculated, and statistical significance was assigned to differences at the 95% confidence level.[16]

Since the focus of our interest was coronary disease, age-adjusted death rates were computed only for those aged 35 and older. Rates were based on 10-year averages over the period 1935 to 1985, using the 5 years on either side of an official census count, with the census figures as the denominator. Although the age composition of the two populations changed dramatically over time, standardization of the rates takes this into account. (See the Appendix for complete census data.)

Results

Table 1 presents age-adjusted mortality rates from myocardial infarction, congestive heart failure, the two combined, and for all causes for the two communities over the course of five decades. The mortality rate for myocardial infarction among Roseto men and women was initially very low but showed a progres-

TABLE 2—Age-Specific Death Rates Per 1000 from Myocardial Infarction, Roseto Men and Women, 1955 to 1974

	1955–1964				1965–1974			
	Bangor		Roseto		Bangor		Roseto	
	Death Rate	No. of Deaths	Death Rate	No. of Deaths	Death Rate	No. of Deaths	Death Rate	No. of Deaths
Men								
35–44	5.3	2	8.3	1	4.0	1	33.3	3
45–54	32.6	14	9.8	1	57.0	20	55.0	6
55–64	117.3	36	29.0	2	95.2	34	92.8	9
65+	216.4	66	211.3	15	248.5	82	268.7	18
Women								
35–44	6.3	3	0		0		0	
45–54	6.6	3	0		0		0	
55–64	43.1	15	15.9	1	33.9	15	15.9	2
65+	140.7	57	120.0	9	146.3	73	181.8	16

TABLE 3—Standardized Mortality Ratios (SMR) of Roseto to Bangor and Confidence Intervals (CI), 1935 to 1984

	1935–1944		1945–1954		1955–1964		1965–1974		1975–1984	
	SMR	CI	SMR	CI	SMR	CI	SMR	CI	SMR	CI
Myocardial infarction										
Men	.58	(.41, 0.75*)	.66	(.52, 0.81*)	.66	(.52, 0.80*)	1.15	(.95, 1.34)	1.03	(.84, 1.22)
Women	.53	(.32, 0.74*)	.70	(.46, 0.93*)	.65	(.46, 0.84*)	1.07	(.78, 1.36)	1.06	(.78, 1.37*)
Congestive heart failure										
Men	.81	(.63, 1.00)	.84	(.61, 1.07)	.88	(.61, 1.15)	.85	(.61, 1.09)	.51	(.27, 0.75*)
Women	1.05	(.82, 1.29)	1.35	(.99, 1.72)	.79	(.54, 1.04)	1.11	(.76, 1.47)	.88	(.50, 1.25)
Myocardial infarction and congestive heart failure										
Men	.71	(.58, 0.84*)	.73	(.60, 0.85*)	.72	(.60, 0.84*)	1.06	(.90, 1.21)	.93	(.78, 1.09)
Women	.89	(.72. 1.06)	1.05	(.83, 1.26)	.71	(.56, 0.86*)	1.09	(.86, 1.31)	1.00	(.78, 1.23)
All causes										
Men	.77	(.69, 0.86*)	.86	(.76, 0.96*)	.80	(.71, 0.89*)	1.08	(.97, 1.19)	.78	(.69, 0.86*)
Women	.96	(.85, 1.06)	1.13	(.98, 1.27)	.79	(.68, 0.90*)	1.17	(1.01, 1.33*)	.76	(.64, 0.87*)

*Significant at confidence levels of 95%; under 1.0 the Roseto rate is lower.

sive rise over the 30-year period from 1935 to 1964. The rate for men in Bangor also rose during this time, while rates for Bangor women peaked in the period from 1955 to 1964. In the decade 1965 to 1974 there was a sharp increase in mortality from myocardial infarction among Roseto men and women. For women, the increase from the 1955 to 1964 decade was entirely in the 65 and older population, but for men it occurred in every age group and most dramatically in all 10-year age groups younger than 65 (see Table 2). In contrast, rates for congestive heart failure generally declined over the course of the five decades. Total mortality rates declined in both communities, although the drop is considerably greater for women than for men.

Table 3 shows the Roseto to Bangor ratios and confidence intervals of mortality rates for total deaths and for myocardial infarction, congestive heart failure, and the two combined. The results indicate that the rates of death from all causes for men are significantly lower in Roseto for all decades except 1965 to 1974. For women they are significantly lower in 1955 to 1964 and 1975 to 1984 but significantly higher from 1965 to 1974.

Although there were no significant differences between the two communities in congestive heart failure until the last decade, the mortality rates for myocardial infarction in both men and women in Roseto were significantly lower than those for Bangor men and women over the initial period of thirty years. After 1965 this difference disappeared. When myocardial infarction and congestive heart failure are combined, Roseto men had a mortality advantage over Bangor men from 1935 until 1965.

Discussion and Conclusion

After a very thorough search of all sources of data for mortality in two small Pennsylvania communities over the course of 50 years, our examination of death certificates has confirmed the earlier inference, based on a shorter span of years, that the death rate from myocardial infarction was lower in Roseto than in immediately adjacent Bangor in three decades prior to 1965. The difference between the two communities is statistically significant despite the small number of myocardial infarctions. The sharp rise that followed involved mainly young Rosetan men and elderly women at a time when the predicted decrease in social cohesion became clearly manifest, as described in detail elsewhere.[14,15] The increase in myocardial in-

farction in Roseto was accompanied by a similar increase in total mortality.

This remarkable pattern suggests systematic differences between the two neighboring communities over the course of at least 30 years—years for which there are many indicators of greater social solidarity and homogeneity in Roseto and no evidence of differences in coronary risk factors.[15] The social changes that occurred in Roseto in the 1960s are reflected in sharply increased rates of heart attack among men under the age of 65. □

Acknowledgments

This research was funded by the National Institute on Aging Grant 1R01AG04957, the Eleanor Naylor Dana Charitable Trust, and The Pew Charitable Trusts.

The authors thank Donald T. Campbell and the many people involved in collecting data for this project for their assistance.

References

1. Stout C, Morrow J, Brandt EN, Wolf S. Study of an Italian-American community in PA; unusually low incidence of death from myocardial infarction. *JAMA.* 1964;188:845.
2. Bruhn JG, Wolf S. *The Roseto Story: An Anatomy of Health.* Norman, Okla: University of Oklahoma Press; 1979.
3. Bruhn JG, Chandler B, Miller C, Wolf S. Social aspects of coronary heart disease in two adjacent ethnically different communities. *Am J Public Health.* 1966;56:1493–1506.
4. Lynn TN, Duncan R, Naughton J, et al. Prevalence of evidence of prior myocardial infarction, hypertension, diabetes and obesity in three neighboring communities in Pennsylvania. *Am J Med Sci.* 1967;254:385–391.
5. Bruhn JG, Phillips B, Wolf S. Social readjustment and illness patterns. Comparison between first, second and third generation Italian-Americans living in the same community. *J Psychosom Med.* 1972;16:387–394.
6. Kiritz S, Moos RH. Review article: physiological effects of social environments. *Psychosom Med.* 1974;36:96–114.
7. Lynch JJ. *The Broken Heart: The Medical Consequences of Loneliness.* New York, NY: Basic Books; 1979.
8. Spittle B, James B. Psychosocial factors and myocardial infarction. *Aust N Z J Psychiatry.* 1977;37:96–114.
9. Jenkins CD. Behavioral risk factors in coronary artery disease. *Heart Disease Behav.* 1978;29:543–562.
10. Rahe RH, Arthur RJ. Life change and illness studies: past history and future directions. *J Hum Stress.* 1978;4(1):3–15.
11. Thoreson RW, Ackerman M. Research review. Cardiac rehabilitation: basic principles and psychosocial factors. *Rehabil Counsel Bull.* 1981;24:223–255.
12. Schwab JJ. Stress from a psychiatric epidemiological perspective. *Stress Med.* 1986;2:211–220.
13. Wolf S, Grace K, Bruhn JG, Stout C. Roseto revisited: further data on the incidence of myocardial infarction in Roseto and neighboring Pennsylvania communities. *Trans Am Clin Climatol Assoc.* 1974;85:100–108.
14. Wolf S, Herrenkohl RC, Lasker J, Egolf B, Phillips B, Bruhn JG. Roseto, Pennsylvania 25 years later—highlights of a medical and sociological review. *Trans Am Clin Climatol Assoc.* 1988;100:57–67.
15. Wolf S, Bruhn JG. *The Power of Clan; A 25-Year Prospective Study of Roseto, Pennsylvania.* New Brunswick, NJ: Transaction Publishers; in press.
16. Dever GE, Champagne F. *Epidemiology in Health Services Management.* Rockville, Md: Aspen; 1984.
17. Census of population. Washington, DC: US Bureau of the Census; 1940, 1950, 1960, 1970, 1980.

APPENDIX—Bangor and Roseto Population Census by Age and Sex

Age Group	1940 M	1940 F	1950 M	1950 F	1960 M	1960 F	1970 M	1970 F	1980 M	1980 F
Bangor										
Under 35	1574	1628	1483	1559	1326	1330	1256	1214	1251	1175
35–44	383	394	478	520	380	477	252	285	208	240
45–54	345	374	383	386	430	458	351	438	215	281
55–64	263	272	291	349	307	348	357	443	298	384
65+	225	229	273	328	305	405	330	499	347	607
Total	2790	2897	2908	3142	2748	3018	2546	2879	2319	2687
Total of both sexes	5687		6050		5766		5425		5006	
Roseto										
Under 35	605	655	485	506	417	424	358	353	366	302
35–44	99	92	125	165	121	143	90	111	57	77
45–54	60	64	83	79	102	145	109	139	93	105
55–64	58	60	44	65	69	63	95	126	90	121
65+	45	40	67	57	71	75	67	88	105	168
Total	867	911	804	872	780	850	719	817	711	773
Total of both sexes	1778		1676		1630		1536		1484	

The Quality of Life in the Year before Death

ABSTRACT

Objectives. Most Americans wish to live a long healthy life, but fear disease and dependency in their last years. Until recently, little has been known about the prevalence of opposite extremes of health in old age, particularly in the period leading up to death.

Methods. We used results from the 1986 National Mortality Followback Survey to estimate proportions of elderly decedents who were "fully functional" or "severely restricted" in the last year of life. Estimates were based on responses from proxies to questions regarding the decedent's functional status, mental awareness, and time spent in institutions.

Results. Approximately 14% of all decedents aged 65 years and older were defined as fully functional in the last year of life; 10% were defined as severely restricted. Proportions varied with the decedent's age and sex, the underlying cause of death, and the presence of other preexisting conditions.

Conclusions. Results from this survey and future surveys can be used to learn more about "successful agers"—their medical histories, their life-styles, and whether their relative number is increasing or decreasing over time. (*Am J Public Health.* 1992; 82:1093–1098)

Harold R. Lentzner, PhD, Elsie R. Pamuk, PhD, Elaine P. Rhodenhiser, Richard Rothenberg, MD, MPH, and Eve Powell-Griner, PhD

Introduction

Americans born in the late 1980s could expect to live into their 70s, and those reaching their 65th birthday at that same time could expect to live almost two additional decades.[1] With the expectation of long life after age 65 has come the desire for accompanying good health, and, correspondingly, the fear of spending one's final years in a state of severe disability and dependency. As was noted in the now decade-old *Healthy People: the Surgeon General's Report on Health Promotion and Disease Prevention,* "older Americans hope for a state of well-being which would allow them to perform at their highest functional capacity. . . . Their greatest fear is of being useless, sick or unable to care for themselves."[2]

Until recently, little has been known about the distribution of functionality in populations of older persons. Several recent studies have looked at predictors of health status among samples of the elderly[3–8]; however, none of these efforts has focused on the period preceding death. This period is important, in part, because those who live long lives and are vibrant until shortly before death may provide the best possible example of "successful aging."[9]

In this study, we developed a simple multidimensional index to estimate extremes in health status in the last 12 months of life among a nationally representative sample of persons who died in 1986. We also examined several demographic, health, and social factors to identify important correlates of being fully functional or, conversely, severely disabled in these final months.

Methods

National Mortality Followback Survey Sample

For this analysis, we extracted all deaths of individuals aged 65 years and older from the 1986 National Mortality Followback Survey (NMFS) (n = 7586). The NMFS is a stratified random sample of approximately 1% of all deaths among US residents (excluding residents of Oregon, who were not sampled) aged 25 years and older that occurred in 1986 (n = 18 733).[10] This survey, conducted by the National Center for Health Statistics, collected detailed information on health status and health care provided in the last year of life and on aspects of lifestyle that are potential risk factors for disease and disability. Most information was collected from proxy respondents by mail questionnaire, by telephone, or by personal interview. Information on the underlying cause of death and all associated causes came from the official death certificate. Because information for comorbid conditions (conditions that occurred or were prevalent at least 12 months before death and were not

Harold R. Lentzner, Elsie R. Pamuk, and Richard Rothenberg are with the National Center for Chronic Disease Prevention and Health Promotion, and Elaine P. Rhodenhiser is with the National Center for Environmental Health and Injury Control, all at the Centers for Disease Control, Atlanta, Ga. Eve Powell-Griner is with the Department of Public Health Sciences, School of Public Health, University of Alabama at Birmingham, Birmingham, Ala.

Requests for reprints should be sent to Harold R. Lentzner, PhD, 1600 Clifton Road, Mail Stop (K30), Atlanta, Ga.

This paper was submitted to the Journal March 6, 1991, and accepted with revisions November 12, 1991.

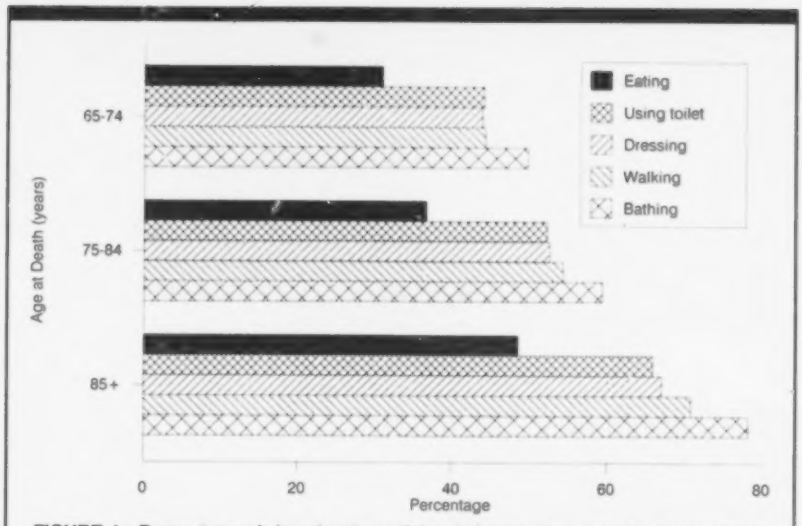


FIGURE 1—Percentage of decedents receiving help or using special equipment to perform activities of daily living in last year of life, by age at death.

the underlying cause of death) was unavailable for a number of decedents, we had a reduced sample (n = 5582) for this portion of the study. However, the characteristics of this subsample closely matched those of the full sample.

Creation of an Index

We examined a subset of responses to questions asked on the NMFS. Proxy respondents were asked if the decedent had "received help or used special equipment" in the last year of life in five separate activities of daily living (ADLs)—eating, using the toilet, dressing, walking, and bathing—and if so, over what period of time. They were asked how often in that same period the decedent had trouble understanding where he or she was, remembering what year it was, or recognizing family or good friends. The respondent was also asked how many nights the decedent spent in hospitals and nursing homes in the last year of life.

Results from all three sets of questions were used to construct a simple additive index of health, and two groups of decedents with opposite profiles were identified. We defined as "fully functional" those who during the last year of life needed no assistance or special equipment to carry out the five basic ADLs listed above, who were lucid until the last few hours or days before death, and who spent less than a week in a hospital or nursing home. At the other extreme were "severely restricted" individuals, who needed help or special equipment for 6 months or more in at least three of five ADLs and more limited assistance in the remaining two, who exhibited at least some decline in mental functioning, and who spent at least half of their last year in a hospital or nursing home.

Statistical Methods

The number and proportion of decedents who were fully functional in their last year and of those who were severely restricted were estimated for three age groups: 65 to 74 years, 75 to 84 years, and 85 years and older. Variances and confidence intervals were computed by using generalized parameters provided by the National Center for Health Statistics.[11]

We used separate logistic regression models to estimate the likelihood of being fully functional or severely restricted in the last year. Decedents who fit these definitions were compared with all other decedents; that is, the fully functional were compared with all decedents experiencing greater restrictions, and the severely restricted were compared with all decedents experiencing fewer restrictions.

We calculated regression coefficients, odds ratios, and variances by using the raw sample reweighted to account for the stratified sampling design but not for the effects of clustering. For both sets of models, the odds ratios reflect the relative likelihood of being fully functional or severely restricted, adjusted for single year of age, and in the presence of other important factors.

Results

Characteristics of the Respondents

The relationship of the respondents to the deceased varied according to the characteristics of the decedent. Although a member of the immediate family was most often the proxy respondent, other relatives, friends, and acquaintances were more likely to be proxies for female than for male decedents (24% vs 15%), for Black than for White decedents (25% vs 19%), and for those who died at age 85 years or older (27%) than for those who died at ages 75 to 84 years (20%) or 65 to 74 years (12%).

Components of the Index

The proportion of respondents requiring assistance with ADLs rose with age at death (Figure 1). For all age groups, decedents were least likely to need help eating and most likely to need help bathing.

Among those who died between the ages of 65 and 74, roughly 15% had had trouble knowing where they were, 13% remembering the year, and about 10% recognizing family or good friends (Figure 2). These proportions increased with age; the oldest decedents were more than twice as likely as the youngest to have cognitive limitations.

Of those who died between the ages of 65 and 74, 7.2% spent more than half of their last year in a nursing home. Among those who died between the ages of 75 and 84, the corresponding figure was 17.4%. But of those who died at the age of 85 years or older, 39.3% spent at least half of their last year in a health care facility.

Health Status and Demographic Characteristics

Approximately 14% of all decedents who died at age 65 years and older were defined as fully functional in the last 12 months of life. A slightly lower proportion, 10%, met our definition of severely restricted. As might be expected, the proportion of decedents at the definitional extremes varied with age in very different ways. One fifth of those who died between the ages of 65 and 74 met our definition of fully functional in the last 12 months, but the percentage dropped to 6% among the oldest group; by contrast, only 3% of the youngest group but 22% of those aged 85 years and older were severely restricted.

Differences by sex were pronounced. A smaller proportion of women than men were fully functional in the last 12 months

of their lives, and a larger proportion were severely restricted. This was true for decedents in all three age groups (Figure 3). For the oldest decedents, the differences are substantial: roughly 10% of the men vs 4% of the women remained unrestricted in the last year, whereas 13% of the men vs 26% of the women met the definition of severely restricted.

We also examined health status by race and marital status at death. We found few significant differences and noted no consistent pattern across marital status subgroups. Of those who died before the age of 85, Whites appeared more likely than Blacks to be either fully functional or severely restricted, although the differences were not always statistically significant.

Health Status and Cause of Death

In general, a higher proportion of those who died of acute myocardial infarction (AMI) were fully functional, and a lower proportion were severely restricted, than were those who died of other causes (other forms of heart disease, cancer, cerebrovascular disease and all other causes) (Table 1). In contrast, fewer than 1 in 20 persons who died of cancer could be defined as fully functional, regardless of age at death. However, a relatively small proportion of these decedents could be categorized—at the other extreme—as severely restricted.

For most causes, a significantly higher proportion of women than men at the older ages were severely restricted in the last year of life, but the sex-specific relationships between cause and full functionality were less uniform. For those dying of heart disease and of all causes other than heart disease, cancer, and stroke, women were less likely to be fully functional; for those who died of cancer, there were no differences. Among persons dying of cerebrovascular diseases, women appeared more likely to be fully functional, although the differences were not large enough to reach statistical significance.

Health Status and Preexisting Comorbid Conditions

At all ages, decedents with a history of stroke were less likely than most others to be fully functional and more likely to be severely restricted in the last year (Table 2). However, decedents with conditions other than stroke were not consistently less likely to be fully functional or more likely to be severely restricted than were decedents with none of these chronic conditions. In most cases, lower proportions of women than men were fully functional and higher proportions were severely restricted for similar types of preexisting conditions.

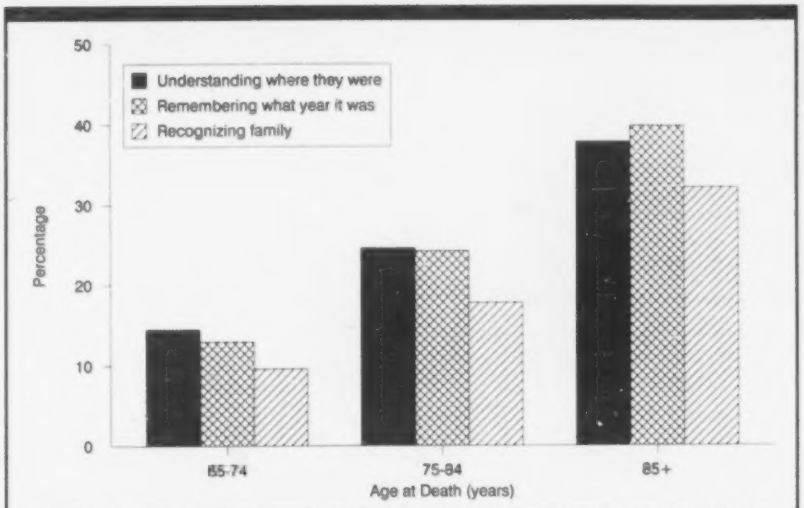


FIGURE 2—Percentage of decedents having cognitive difficulties in last year of life, by age at death.

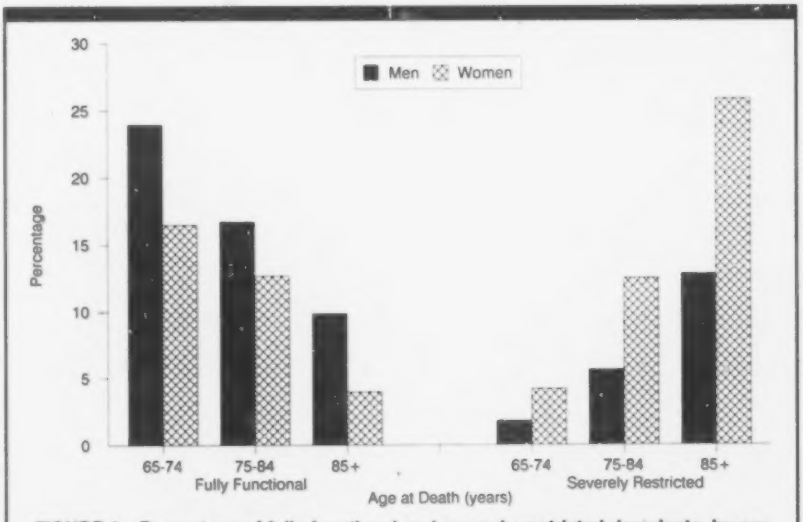


FIGURE 3—Percentage of fully functional and severely restricted decedents, by sex and age at death.

Multivariate Analysis

Sex was highly predictive of health status even after we controlled for age, race, marital status, cause of death, and six preexisting conditions. In the last year of life, women were about 40% less likely than men to have been fully functional and 70% more likely to have been severely restricted (Table 3).

Decedents of races other than White were less likely to be either fully functional or severely restricted, although the odds ratio for these races was significantly different only for the model that examined severe restriction. Marital status at time of death did not appear to be related to full functionality. In contrast, being other than married—widowed, divorced, separated, or never married—doubled the odds of being severely restricted in the last year of life.

The underlying cause of death was associated with the likelihood of being fully functional in the year before death. Those who died from AMI were far more likely to have been fully functional than those who died from any other cause. Decedents with a history of stroke or chronic

TABLE 1—Percentages of Fully Functional and Severely Restricted Decedents, by Age at Death, Sex, and Cause of Death

Cause of Death[a]	Age at Death, y					
	65–74		75–84		85+	
	Men	Women	Men	Women	Men	Women
Fully Functional						
Heart disease						
AMI	48.6	37.4	36.1	25.6	24.4	10.8
All other	31.5	23.0	26.1	16.8	11.3	3.3
Cerebrovascular	19.1	28.1	10.5	13.4	1.9	4.9
Cancer	4.4	3.2	5.4	4.2	4.2	2.2
Other	23.1	13.8	8.0	7.8	7.7	2.5
Severely Restricted						
Heart disease						
AMI	0.5	0.0	2.4	3.2	2.9	17.8
All other	2.3	5.4	4.7	15.0	15.7	25.5
Cerebrovascular	5.0	7.6	12.7	25.3	11.4	27.4
Cancer	1.2	1.1	1.8	4.4	7.4	11.7
Other	2.1	8.3	8.9	15.4	15.7	32.8

[a]Causes of death are coded according to the *International Classification of Diseases*, Ninth Revision:
Heart disease
Acute myocardial infarction (AMI), 410
All other, 390–398, 402, 404, 411–429
Cerebrovascular, 430–438
Cancer, 140–208
Other, all other codes.

TABLE 2—Percentages of Fully Functional and Severely Restricted Decedents, by Age at Death, Sex, and Type of Preexisting Condition

Preexisting Condition[a]	Age at Death, y					
	65–74		75–84		85+	
	Men	Women	Men	Women	Men	Women
Fully Functional						
Hypertension	28.6	17.8	15.4	12.9	7.9	4.4
Heart disease	27.4	18.3	17.7	14.8	10.8	4.0
Cancer	38.0	9.3	12.7	15.9	9.3	0.0
Stroke	15.2	5.8	2.2	4.5	1.8	1.6
Lung condition	19.2	11.8	13.2	9.1	4.8	3.3
Diabetes	20.8	18.1	11.3	12.2	12.1	1.4
None	23.2	16.6	19.3	11.0	13.1	5.6
Severely Restricted						
Hypertension	2.5	4.6	5.3	12.3	12.3	21.0
Heart disease	0.4	4.7	5.5	9.2	9.2	21.7
Cancer	0.0	0.0	4.2	13.8	24.8	22.8
Stroke	4.2	13.3	17.5	23.4	30.8	39.1
Lung condition	0.9	6.4	5.8	9.9	4.9	23.0
Diabetes	2.8	8.1	7.4	14.0	20.2	24.2
None	0.5	4.7	4.0	14.0	9.7	26.6

[a]A preexisting condition is defined as a condition that existed 12 months prior to death and that was not the underlying cause of death. Categories are not mutually exclusive.

obstructive pulmonary disease, or with a history of two or more preexisting conditions, were less likely to have been fully functional than decedents with none of these conditions. Dying from heart disease other than AMI, dying from stroke or "other causes," and having a history of stroke or of three or more preexisting conditions were all associated with a greater likelihood of very poor functional status than the appropriate reference categories.

We used sex-specific models to evaluate the relative importance of each of the remaining variables for men and women. In general, the direction and strength of the associations were comparable to the first set of models. Few differences were apparent in the fully functional models, except that preexisting conditions seemed to be more important for men. Marital status appeared to be more strongly associated with the likelihood of having been severely restricted in women than in men, whereas cause of death and disease history were more important for men than for women. Women decedents with hypertension and no other preexisting condition were less likely to have been severely restricted than were women decedents with no preexisting conditions.

Discussion

Our results suggest that elderly decedents are fully functional or severely restricted in the 12 months preceding their death in roughly equal proportions, but that the proportion varies with age. Using provisional mortality counts for 1990 (the most recent year for which data are available) and applying the age- and sex-specific prevalences for functionality, we would estimate that about 216 000 elderly decedents experienced little decline in physical independence or mental awareness and few nights in health care facilities during the last year of life.[12] On the other hand, we estimate that 169 000 elderly decedents suffered both mental and physical impairment and were institutionalized for the majority of their last 12 months of life. However, when we excluded the decedents who failed to reach average life expectancy (whose deaths must be regarded as premature no matter what the circumstances), we found that a much larger proportion of decedents required considerable care. Of the oldest decedents—those aged 85 years and older at the time of death—we estimate that about 100 000 died with profound physical impairments and required assistance through much of the last year of life, whereas only about 28 000 were fully functional.

As the oldest-aged population grows in the United States and in other aging societies, the number who are severely disabled may be expected to grow, even with declining mortality rates, unless the present pattern changes. By the year 2000, the number of persons aged 85 years and older who will be severely restricted in their final months may be expected to increase by 50%, given recent official population projections.[13] If health care costs

continue to escalate at present rates, the costs of the care associated with their disabilities may be expected to increase by a much larger percentage.

Our results showing sex as a significant predictor of disability are consistent with findings from recent research on predictors of "healthy aging."[4,5] Guralnik and Kaplan note that although men tend to live shorter lives than women, they are less likely to be functionally impaired.[5] Unfortunately, we cannot determine whether other diseases are responsible for differences by sex in the quality of the last year of life without a more complete inventory of prevalent chronic diseases than is presently on the NMFS. For example, arthritis is a disease that frequently ends in severe disability, and arthritic conditions, whether self-reported or physician-diagnosed, are more common among older women than among older men.[14,15] After controlling for 13 disabling conditions (including arthritis), Verbrugge et al. found that women aged 55 years and older had only slightly higher levels of disability than did men in that age group.[3]

In addition, our definition of "severely restricted" may be responsible for some of the apparent male advantage in functionality, because men are more likely to have a living spouse than are women and are thus more likely to avoid institutionalization in the last year of life. However, when we examined the separate components of the index, we found that female decedents, besides having poorer results than male decedents for ADLs and institutional use, were also more likely to have had cognitive impairments.

Our results also show that factors other than sex are associated with quality of life in the last year. Dying of anything other than a heart attack sharply decreased the likelihood of being fully functional in the 12 months prior to death. Moreover, death from diseases or conditions with prolonged morbid processes that are associated with both physical and mental deterioration, such as stroke and diseases of the heart other than AMI, increased the likelihood that decedents experienced restriction and dependency in their last year of life. Our research suggests that, all other things being equal, the consequence of replacing an earlier death from AMI with a later death from stroke is a greater likelihood of functional disability. This finding suggests that changes in the cause-of-death structure will have an impact on the aggregate assessment of the quality of life in the year before death.

TABLE 3—Age-Controlled Odds Ratios from Multivariate Logistic Models

	Men		Women		Total	
	FF	SR	FF	SR	FF	SR
Sex						
Male (reference)	...	...	...	...	1.00	1.00
Female	...	...	...	...	.60*	1.73*
Race						
White (reference)	1.00	1.00	1.00	1.00	1.00	1.00
All other	.83	.53	.81	.55*	.83	.54*
Marital Status						
Married (reference)	1.00	1.00	1.00	1.00	1.00	1.00
Widowed/divorced/ separated/never married	1.13	1.88*	.85	2.43*	1.01	2.07*
Cause of death						
AMI (reference)	1.00	1.00	1.00	1.00	1.00	1.00
Other heart disease	.51*	3.69*	.39*	2.28*	.47*	2.56*
Stroke	.17*	4.72*	.36*	3.30*	.27*	3.64*
Cancer	.05*	1.95	.06*	.65	.06*	.89
Other	.20*	4.91*	.17*	2.97*	.20*	3.31*
Preexisting conditions						
None (reference)	1.00	1.00	1.00	1.00	1.00	1.00
Single condition						
Hypertension	1.19	1.04	1.09	.55*	1.08	.64*
Heart disease	.93	.93	.58	.68	.81	.72
Cancer	1.12	1.94	.82	.87	.95	.90
Stroke	.02*	6.85*	...	1.15	.02*	2.33*
COPD	.45*	1.47	1.31	.92	.62*	1.06
Diabetes	.66	2.40*	.88	.72	.72	1.08
Two conditions	.58*	1.86*	.76	.98	.62*	1.16
Three or more conditions	.32*	2.30*	.41*	1.35	.34*	1.56*

Note. FF = fully functional; SR = severely restricted; AMI = acute myocardial infarction; COPD = chronic obstructive pulmonary disease.
[a]Stroke was omitted from the model for fully functional women.
*$P<0.5$.

The validity of proxy response will influence the results. Responses from proxies are believed to be most valid when they are based on directly visible phenomena—for example, functional activities, or habits or events such as hospitalizations.[16] The amount of time the proxy spends assisting the patient also influences the proxy's assessment of functional impairment.[17]

To minimize these potential problems of proxy response, we focused on those decedents who were either so healthy or so sick in the last year of life that their condition was a clear matter of record. We evaluated the sensitivity of the results to changes in the definitional criteria by including adjacent categories in the analysis. These changes in definitions resulted in larger numbers of decedents in both the fully functional and the severely restricted groups and, as expected, somewhat lessened the distinction between the two. However, the distinction's persistence suggested that definitional alteration would produce only minor changes in the results.

The NMFS obtained information on the last year of life only. Although it is unlikely that persons who were fully functional in their last year were partially or severely restricted in the years preceding, the reverse may be true: restriction in the final year may follow a long period of functionality. Even though the picture derived from NMFS data is incomplete, we can begin to develop profiles of "successful agers"—men and women who lived into their ninth decade and were fully functional through most of their last year of life—as well as profiles of those who are not so fortunate. We know, for example, that only 25% of the women and 40% of the men who were successful agers had any of the comorbid conditions identified in the NMFS. By contrast, of those who died early with severe restrictions, 66% of the women and more than 90% of the men had at least one of the preexisting conditions. Moreover, persons who died before age 80 after a period of severe restriction were more likely to have died of causes other than AMI than were successful agers.

A valuable addition to a health profile of successful agers would be the quantification of the benefits, if any, obtained from a lifetime of healthy living or risk avoidance. The 1986 NMFS requested information from proxies on the decedent's lifetime smoking, drinking, dietary, and exercise habits and on the decedent's weight. Concerns about data quality and response distribution made it difficult to evaluate the association between successful aging and a multidimensional healthy living index. However, we found that successful agers (defined here as persons who died at age 80 or older and were fully functional) of either sex were less likely than all others to have ever smoked, although the results were significant for men only (data not shown). There was also some evidence that successful agers were less likely than others either to have abused alcohol or to have abstained from it, and that women who were successful agers were less likely than others to have been severely overweight or underweight.

These results are only suggestive. A more fundamental question that cannot be answered here is whether the life-style changes that improve the probability of survival also improve the probability of remaining healthy. Future research should be directed toward answering this important question.

The concept of successful aging should embody both longevity and vitality. That we are experiencing unprecedented gains in life expectancy is well recognized. The quality of life in the 12 months before death provides a potentially important measure of the success or failure of those gains. The next National Mortality Followback Survey, to be conducted in 1993, will provide an important second point for a vector of the quality of aging. □

References

1. *Health, United States, 1989.* Hyattsville, Md: National Center for Health Statistics; 1990.
2. *Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention.* Rockville, Md: US Dept of Health, Education, and Welfare, Public Health Service, Office of the Assistant Secretary for Health and the Surgeon General; 1979, US Dept of Health, Education, and Welfare publication PHS 79-55071.
3. Verbrugge L, Lepkowski JM, Imanaka Y. Comorbidity and its impact on disability. *Milbank Q.* 1990;67:450–484.
4. Harris T, Kovar MG, Suzman R, Kleinman JC, Feldman JJ. Longitudinal study of physical ability in the oldest-old. *Am J Public Health.* 1989;79:698–702.
5. Guralnik JM, Kaplan GA. Predictors of healthy aging: prospective evidence from the Alameda County Study. *Am J Public Health.* 1989;79:703–708.
6. Guralnik JM, LaCroix AZ, Branch LG, Kasl SV, Wallace RB. Morbidity and disability in older persons in the years prior to death. *Am J Public Health.* 1991;81:443–447.
7. Guralnik JM. Prospects for the compression of morbidity: The challenge posed by increasing disability in the years prior to death. *J Aging Health.* 1991;3:138–154.
8. Roos NP, Montgomery P, Roos LL. Health care utilization in the years prior to death. *Milbank Q.* 1987;65:231–254.
9. Fries JF. Aging, natural death, and the compression of morbidity. *N Engl J Med.* 1980;303:130–135.
10. Seeman I, Poe G, Powell-Griner E. Development methods and response characteristics: 1986 National Mortality Followback Survey. *Vital Health Stat [2].* In press.
11. Kapantais G, Powell-Griner E. Characteristics of Persons Dying from AIDS: Preliminary Data from the 1986 National Mortality Followback Survey. *Advance data from vital and health statistics,* no. 173. Hyattsville, Md: National Center for Health Statistics; 1989.
12. Annual summary of births, marriages, divorces, and deaths: United States, 1990. *Monthly vital statistics report, vol. 39, no. 13.* Hyattsville, Md: National Center for Health Statistics; 1991.
13. Wade A. *Social Security Area Population Projections: 1988.* Baltimore, Md: Social Security Administration; June 1988.
14. Verbrugge L. Arthritis and disability in older adults. Presented at the annual meeting of the Pc[illegible]n Association of America; April 21[illegible]8; New Orleans, La.
15. Branch LG, Meyers AR. Assessing physical function in the elderly. *Clin Geriatr Med.* 1987;3:29–51.
16. Epstein AM, Hall JA, Tognetti J, Son LH, Conant L. Using proxies to evaluate quality of life. *Med Care.* 1989;27:S91–S98.
17. Rodgers WL, Herzog AR. The consequences of accepting proxy respondents on total survey error for elderly populations. In: Fowler FJ, ed. *Conference Proceedings: Health Survey Research Methods.* Washington, DC: National Center for Health Statistics Research; 1990.

HMO Membership, Treatment, and Mortality Risk among Prostatic Cancer Patients

Howard P. Greenwald, PhD, and Curtis J. Henke, PhD

ABSTRACT

Objectives. Treatment and mortality risk were compared between prostate cancer patients receiving care in fee-for-service settings and those receiving care in a health maintenance organization (HMO).

Methods. Two samples were obtained from a population-based tumor registry. Patients in the first sample (n = 201) were interviewed shortly after diagnosis to obtain data on income, education, overall health status, and expenditures for health care. These data were combined with information from the tumor registry on cancer stage, age, treatment, place of residence, and source of care. Only tumor registry data were obtained for most patients in the second sample (n = 962). For both samples, survival time was monitored for up to 80 months.

Results. Multivariate analysis of data from the interviewed sample indicated that HMO patients were less likely to receive surgery but more likely to receive radiation therapy than were those in fee-for-service settings. Mortality risk was lower for the HMO patients than for those in fee-for-service plans. Findings based on the second sample were nearly identical.

Conclusions. This study suggests that HMOs may offer important advantages to lower-income patients at risk for specific life-threatening diseases. (*Am J Public Health.* 1992;82: 1099–1104)

Introduction

Government, business, and private citizens alike look increasingly to the health maintenance organization (HMO) as a source of affordable, high-quality health care. Supporters of HMOs argue that these organizations deliver care more economically than the fee-for-service system[1–5] and provide care of a quality acceptable to a wide variety of individuals.[6] Several commentators and researchers, however, have suggested that HMOs reduce the quality of care available to at least some consumers.[7,8] One major study reported that low-income HMO members in poor health experience more serious symptoms and greater mortality risk than do those in a "free" fee-for-service plan,[9] although a later analysis indicated no differences.[10] A study of cancer patients reported that HMO members wait longer between diagnosis and the beginning of treatment than do those in fee-for-service.[11]

In this study we examined the treatment and survival of prostatic cancer patients in a large, well-established HMO and in fee-for-service settings in the surrounding community. We compared (1) stage at detection, (2) primary cancer treatment received during the 6 months following initial diagnosis, and (3) mortality risk among patients in the HMO and fee-for-service plans. These areas are directly pertinent to the quality of care. Stage at detection and treatment are important aspects of the process of care; mortality (or its avoidance) is perhaps the key dimension of outcome.

High-quality care of patients with prostatic cancer is an important objective for health professionals. In 1990, the American Cancer Society projected that prostatic cancer would be the most commonly occurring malignancy among males in the United States and that it would account for over 10% of male cancer deaths.[12] The efficacy of many widely used treatment interventions is disputed.[13] However, a recent comprehensive review of cancer mortality indicates that real survival rates, that is, those not explainable merely as statistical artifacts associated with earlier detection, have risen over the past four decades because of improved tumor evaluation, treatment planning, and radiation therapy.[14]

Methods

The investigators analyzed data on two overlapping samples of persons with prostatic cancer. To obtain the first sample (sample 1), they identified and interviewed persons in King and Pierce Counties, Washington (whose major cities are Seattle and Tacoma), who were originally diagnosed with this malignancy between 1980 and 1982. Patients were identified through the Cancer Surveillance System (CSS), a population-based tumor registry maintained by the Fred Hutchinson Cancer Research Center in Seattle. Operating under contract with the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program, the CSS attempts to register all cancer cases that occur in a 13-county area in north-

Howard P. Greenwald is with the Sacramento Center, School of Public Administration, University of Southern California, Sacramento. Curtis J. Henke is with the Department of Medicine, University of California, San Francisco.

Requests for reprints should be sent to Howard P. Greenwald, PhD, Sacramento Center, School of Public Administration, University of Southern California, 1201 J Street, Sacramento, CA 95814-2919.

This paper was submitted to the Journal February 21, 1991, and accepted with revisions December 10, 1991.

western Washington State within 3 months of diagnosis. All individuals identified as HMO members belonged to Group Health Cooperative of Puget Sound, the only major HMO operating in the area in the early 1980s.[15]

Clinical and demographic information was obtained from the CSS, supplemented by face-to-face interviews in sample 1. CSS records provided data on stage, date of diagnosis, location of residence, source of care, diagnostic procedures, initial cancer therapy, and date of death in the event of mortality. Face-to-face interviews were required to obtain data on several patient characteristics capable of affecting the process and outcome of care, such as income, education, direct payment for health services, and general health status. Although previous studies using CSS data exclusively have addressed income and education,[16] they have inferred characteristics of individuals on the basis of their places of residence, a procedure of uncertain accuracy.

The CSS constitutes a highly reliable means of identifying cancer patients and monitoring their survival. An evaluation of the CSS has recently demonstrated that fewer than 2.5% of the cancer cases in its catchment area are missed by the registry.[17] The National Cancer Institute requires the registry to update the status of all living patients within 18 months following diagnosis or last contact.[18] This update is accomplished by contacting the hospitals at which patients were originally treated and the physicians who cared for them[19] and by monitoring the state's computerized file of death certificates.[20]

To select subjects for sample 1, the investigators used a quota sampling procedure with the goal of finding 200 consecutive patients newly diagnosed with prostatic cancer in the pertinent geographical area. Because the interview schedule was long and required significant effort to administer and complete, this sample was restricted to a limited number of patients, all 80 years of age and younger. Human-subjects considerations required the investigators to obtain the physician's permission before his or her patient could be asked for an interview. Physician permission was requested for 377 patients and received for 297. Most refusals (about 75) came from a small number of non-HMO physicians who explained that they were unwilling to allocate staff time or that they disputed the need for a study of pain and quality of life among cancer patients (the initial purpose of the data collection). These physicians refused the investigators access to all of their patients; they did not selectively grant permission to contact some while withholding it for others. The researchers succeeded in contacting 264 patients and obtained sufficient interview data from 201 for inclusion in the study.

For the second sample (sample 2), the investigators selected 1000 consecutive cases of newly diagnosed prostatic cancer from the CSS that were entered between 1981 and 1982, when most of the interviewing of the first sample took place. For sample 2, only data from the CSS records were obtained. Because consent was required from neither the physician nor the patient, data from all 1000 cases were available. (Data from 38 patients who were followed by the CSS for less than 1 month were not included in the analyses.)

To assess bias in sample 1, the investigators compared subjects in sample 2 who were interviewed (and thus also included in sample 1) with subjects in sample 2 who were not interviewed. Among those not interviewed in sample 2, subjects comparable to those in sample 1 were those 80 years of age and younger who had been followed for at least 2 months by the CSS (n = 784). The 2-month criterion was used for this comparison to eliminate those who would have been unavailable for interview because they had died very soon after diagnosis or because diagnosis had been made at autopsy. A total of 174 subjects in sample 2 were also included in sample 1. Those interviewed tended to be younger (mean age: interviewed = 67.5 years, SE = .56; not interviewed = 68.9 years, SE = .29), were more likely to live in an urban census tract (percentage urban: interviewed = 89.7, SE = 2.32; not interviewed = 81.1, SE = 1.58), and were more likely to have been diagnosed at a higher stage (percentage stage II: interviewed = 16.1, SE = 2.93; not interviewed = 12.1, SE = 1.32; percentage stage III: interviewed = 27.6, SE = 3.40; not interviewed = 15.6, SE = 1.47). HMO patients were overrepresented among those interviewed (percentage in the HMO: interviewed = 20.1, SE = 3.08; not interviewed = 6.6, SE = 1.00), probably because no HMO physician refused access to all of his or her patients.

The investigators used ordered logistic regression to assess the relation of HMO membership to stage at detection, independent of the patient's background characteristics (age, income, education, and place of residence). Binary logistic regression was used to predict dichotomous variables indicating whether or not a specific type of treatment—surgery, radiation therapy, chemotherapy, or multimodal treatment (defined as therapy using two or more modalities)—was received. Dichotomous logistic analysis was appropriate here because individual treatment modalities were not mutually exclusive. The Cox proportional hazards model[21] was used to assess the effect of HMO membership on mortality risk independent of individual background characteristics.

Results

Table 1 presents means and standard errors for variables in the analyses that follow. For sample 2, the figures presented in Table 1 and the tables that follow are based on 962 patients. Among the original 1000 patients in sample 2, 38 patients were followed by the CSS for less than 1 month. Because these patients presumably were diagnosed when terminal or at autopsy, the data they generated did not appear to be pertinent to a study of survival and related treatment.

The data shown in Table 1 suggest that cancer stage does not systematically vary across health care source. Ordered logistic regression confirms this initial observation. In this and subsequent tables, cancer stage (stage at diagnosis) is represented as a series of dichotomous variables (stage I = local, II = regional, III = distant spread), coded 0 or 1 as appropriate for each observation. In an equation predicting stage that was based on sample 1 and included age, urban residence, education, and income on the right-hand side, HMO membership had a coefficient of .183 (SE = .352, P = .683). In an equation based on sample 2 that included age and urban residence, HMO membership had a coefficient of .231 (SE = .243, P = .348).

Table 2 summarizes a binary logistic regression analysis of data from sample 1 on key therapies used for prostatic cancer. By statistically significant margins, HMO members were less likely to receive surgery but more likely to receive radiation therapy than were those in fee-for-service plans. No statistically significant relations are visible between HMO membership and hormone or multimodal treatment. In these equations and those presented below, stage is represented by two dichotomous variables (stage II and stage III); the coefficients on these variables reflect differences in the dependent variable from stage I, which is omitted from the equations.

Tables 3 and 4 present Cox proportional hazards models predicting mortality

TABLE 1—Selected Characteristics of Prostatic Cancer Patients by Health Care Source

	Sample 1			Sample 2		
	HMO (n = 44)	Fee for Service (n = 157)	All Sources (n = 201)	HMO (n = 87)	Fee-for-Service (n = 875)	All Sources (n = 962)
Mean age, y	68.1 (1.20)[a]	67.9 (0.57)	68.0 (0.52)	71.2 (0.96)	71.7 (0.32)	71.6 (0.30)
Mean income, $ thousands	28.1 (2.33)	28.3 (1.41)	28.3 (1.20)	...	...	...
Mean education, y	13.1 (0.59)	12.5 (0.29)	12.7 (0.26)	...	...	...
Urban, %	88.4 (4.95)	89.7 (2.45)	89.4 (2.19)	89.7 (3.28)	83.0 (1.13)	83.6 (1.20)
Treatment						
Surgery, %	44.1 (8.64)	75.4 (3.63)	69.3 (3.49)	54.0 (5.37)	72.9 (1.50)	71.2 (1.46)
Hormone, %	20.6 (7.04)	32.1 (3.96)	29.9 (2.48)	23.0 (4.54)	29.4 (1.55)	28.8 (1.47)
Radiation, %	55.9 (8.64)	33.6 (4.01)	37.9 (3.69)	50.6 (5.39)	24.0 (1.45)	26.5 (1.43)
Multimodal, %	20.9 (6.28)	36.9 (3.86)	33.5 (3.35)	31.0 (4.99)	28.2 (1.52)	28.5 (1.46)
Cancer stage						
I, %	46.5 (7.70)	52.2 (4.00)	51.0 (3.54)	52.9 (5.38)	52.8 (1.69)	52.8 (1.61)
II, %	20.9 (6.28)	14.0 (2.78)	15.5 (2.57)	10.3 (3.28)	12.5 (1.11)	12.3 (1.06)
III, %	23.3 (6.52)	26.1 (3.52)	25.5 (3.09)	24.1 (4.61)	17.3 (1.28)	17.9 (1.24)
Survived 5 years, %	76.7 (6.52)	61.8 (3.89)	65.0 (3.38)	69.0 (4.99)	58.1 (1.67)	59.0 (1.57)

[a] Standard errors are in parentheses.

risk based on sample 1. Four variables in Table 3 have statistically significant relationships with mortality risk. HMO membership predicts a lower risk of mortality: individuals in the HMO apparently faced a risk of dying in any given month following diagnosis that was about 38% of that of fee-for-service patients. Educational level predicts mortality risk, each year of education lowering mortality risk by about 9%. As expected, the two variables representing cancer stage are strong predictors of mortality risk.

Figure 1 presents mortality curves for HMO and fee-for-service patients. These curves illustrate relative risks of HMO and fee-for-service patients in each month following diagnosis. Part A of Figure 1 compares the observed proportions of HMO and fee-for-service patients surviving at the beginning of each month. Part B of Figure 1 presents mortality curves based on estimates of risk controlled for disease and background variables using the Cox equation summarized in Table 3.

Table 4 presents a model that assesses the effects of family income level (low, middle, and high) and of HMO membership within specific income categories. For the purposes of this study, individuals reporting family incomes of under $20 000 in the year preceding their interviews are considered to have low incomes; those with incomes between $20 000 and $34 999 are considered to have middle incomes; those with incomes over $35 000 are considered to have high incomes. The equation includes dichotomous variables representing membership in the middle-

TABLE 2—Logistic Regression Analysis of Health Maintenance Organization (HMO) Membership and Cancer Treatment Received, Sample 1

	Treatment Received							
	Surgery		Radiation		Hormone		Multimodal	
HMO member	−1.361**	(0.418)[a]	1.095*	(0.441)	−1.008	(0.632)	−.784	(0.426)
Age	.011	(0.028)	.011	(0.028)	.035	(0.036)	.011	(0.025)
Education	−.127*	(0.054)	.073	(0.051)	−.100	(0.070)	−.072	(0.049)
Income	−.035	(0.062)	.179**	(0.062)	.021	(0.084)	.066	(0.056)
Urban	.711	(0.579)	1.004	(0.730)	−1.288	(0.734)	.544	(0.554)
Cancer stage								
II	−.231	(0.524)	.011	(0.512)	1.074	(0.644)	.306	(0.471)
III	−.543	(0.426)	−1.385**	(0.497)	3.599**	(0.540)	1.129**	(0.372)

[a] Standard errors are in parentheses.
*$P < .05$; **$P < .01$.

TABLE 3—Relation of Health Maintenance Organization (HMO) Membership and Background Variables to Relative Mortality Risk, Sample 1

Patient Characteristic	Coefficient	Risk Ratio	Standard Error	Risk Ratio 95% CI
HMO member	−0.970**	0.379	0.334	0.196, 0.734
Age	−0.013	0.987	0.020	0.949, 1.027
Education	−0.089*	0.915	0.036	0.852, 0.982
Income	−0.063	0.939	0.047	0.856, 1.031
Urban	−0.488	0.614	0.321	0.325, 1.159
Cancer stage				
II	0.866*	2.377	0.358	1.170, 4.830
III	1.676**	5.344	0.271	3.125, 9.139

*$P < .05$; **$P < .01$.

TABLE 4—Relation of Health Maintenance Organization (HMO) Membership among Patients at Specific Income Levels to Relative Mortality Risk, Sample 1

Patient Characteristic	Coefficient	Risk Ratio	Standard Error	Risk Ratio 95% CI
Age	−0.010	0.990	0.021	0.949, 1.032
Education	−0.095**	0.909	0.036	0.847, 0.977
Urban	−0.569	0.566	0.326	0.297, 1.079
Cancer stage				
II	0.632	1.881	0.366	0.911, 3.883
III	1.715**	5.557	0.266	3.282, 9.409
Income				
Middle	−0.528	0.590	0.335	0.304, 1.145
High	−0.682	0.506	0.411	0.224, 1.141
HMO member				
Low income	−1.825**	0.161	0.538	0.056, 0.468
Middle income	0.017	1.017	0.562	0.321, 3.220
High income	0.185	1.203	0.668	0.321, 4.516

*$P < .05$; **$P < .01$.

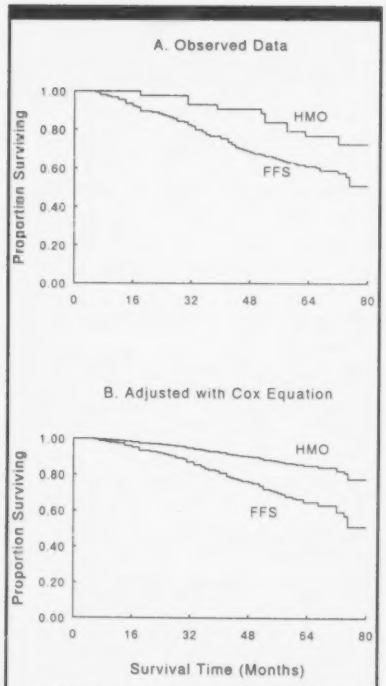


FIGURE 1—Survival by treatment setting. HMO = health maintenance organization; FFS = fee-for-service.

income and high-income categories. Coefficients on these variables represent differences between each income category and low income, a variable omitted from the equation. No significant relationship between either income category and mortality risk is observed.

The table does, however, indicate important income-related differences in the effects of HMO membership on mortality. The equation summarized here includes interaction terms representing HMO membership within the low-, middle-, and high-income categories. Coefficients on these interaction terms indicate the difference in mortality risk of HMO versus fee-for-service patients within each income category. Among low-income patients, mortality risk is lower for HMO members than for those in fee-for-service by a statistically significant margin. No statistically significant differences are observed for patients in the middle- and high-income categories. Figure 2 illustrates these relationships, presenting survival curves computed on the basis of the Cox equations presented in Table 4.

A similar equation including two measures of general health status—an abbreviated (5-point) Karnofsky Scale and the Sickness Impact Profile Scoring Dimension 1[22]—produces results very similar to those presented in Table 4.

To identify patients in "free" fee-for-service settings, the interviewers asked subjects in the face-to-face interviews whether they had made any direct payment for medical care other than insurance premiums in the 12 months before their interviews. A Cox proportional hazards model, similar to the one represented in Table 4, estimated only for individuals reporting no direct payment (n = 88) indicates results very similar to those shown in Table 4. The mortality risk ratio for low-income HMO members versus those in "free" fee-for-service plans was .176 (95% CI = .054, .576), significant at the .01 level. The risk ratio for low-income HMO members versus fee-for-service patients obtained on the basis of a sample segment (n = 157) of HMO members and fee-for-service patients reporting direct payment was .181 (95% CI = .059, .561), $P < .01$.

Many practitioners believe that the implications of prostatic cancer for life expectancy are very different for older and younger patients.[23] But among persons 72 years of age and younger (n = 136), the risk ratio for the low-income HMO members versus fee-for-service patients was .100 (95% CI = .012, .818), $P < .05$. A similar finding was obtained for those aged 65 and younger, although this finding was not statistically significant owing to the small number of observations (n = 67).

Tables 5 and 6 present findings from analyses similar to those reported above but based on data from sample 2. In these tables only age, urban residence, and stage appear as control variables; the CSS does not include information on income, education, or general health status. The results shown in Table 5 are very similar to those shown in Table 2: HMO members were less likely to receive surgery and more likely to receive radiation therapy. HMO members also were less likely to receive hormone therapy than were those in fee-for-service plans. Table 6 presents results very similar to those shown in Table 3: HMO members experience a mortality risk lower than that of fee-for-service patients.

Discussion

In both samples examined in this study, HMO patients were less likely to receive surgery and more likely to receive radiation therapy than were those in fee-for-service plans. In one of the samples analyzed, HMO members also appeared less likely to receive hormone therapy. Mortality risk was lower for the HMO members, a difference that was greatest among those in the lowest income category used in the analysis.

Several features of this study require discussion before conclusions may be drawn. First, the study used a nonexperimental design, admitting the possibility that subjects who selected the HMO, pay fee-for-service, and free fee-for-service plans for their care may have differed in health status, health risks, and behavioral predispositions. Acknowledging this possibility, the investigators included key patient characteristics other than source of care in the analysis of mortality risk. The relationships reported here between

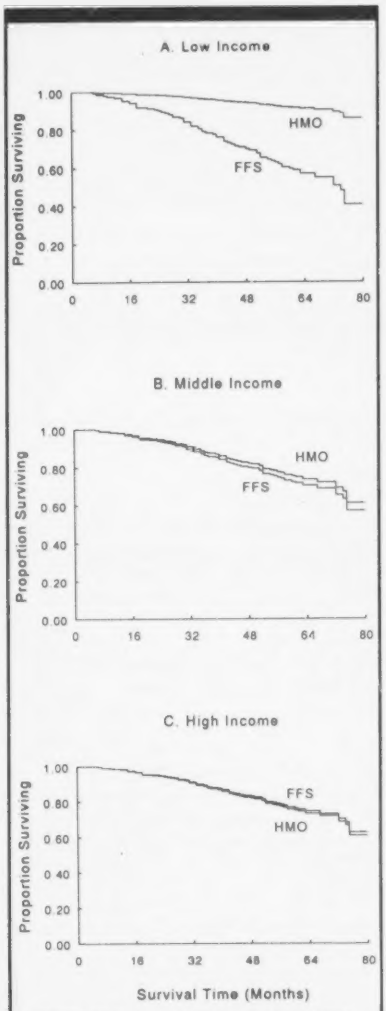


FIGURE 2—Survival by treatment setting by income category. HMO = health maintenance organization; FFS = fee-for-service.

TABLE 5—Logistic Regression Analysis of Health Maintenance Organization (HMO) Membership and Cancer Treatment Received, Sample 2

Patient Characteristic	Treatment Received			
	Surgery	Radiation	Hormone	Multimodal
HMO member	−.799** (0.234)[a]	1.406** (0.252)	−.689* (0.320)	.069 (0.253)
Age	−.012 (0.008)	−.075** (0.009)	.051** (0.009)	−.020* (0.008)
Urban	−.023 (0.200)	.524* (0.238)	.115 (0.236)	.080 (0.202)
Cancer stage				
II	−.286 (0.225)	1.081** (0.223)	1.055** (0.233)	1.048** (0.209)
III	−1.052* (0.179)	−.800** (0.248)	2.942** (0.220)	1.060** (0.161)

[a] Standard errors are in parentheses.
*$P < .05$; **$P < .01$.

TABLE 6—Relation of Health Maintenance Organization (HMO) Membership and Background Variables to Relative Mortality Risk, Sample 2

Variable	Coefficient	Risk Ratio	Standard Error	Risk Ratio 95% CI
HMO member	−0.306*	0.736	0.150	0.547, 0.991
Age	0.069**	1.071	0.005	1.061, 1.082
Urban	−0.326**	0.722	0.112	0.578, 0.901
Cancer stage				
II	0.499**	1.647	0.125	1.286, 2.110
III	1.248**	3.483	0.100	2.858, 4.246

*$P < .05$; **$P < .01$.

HMO membership and mortality risk remain statistically significant even after these dimensions have been held constant according to widely used methods of assessment.

Because cases in sample 1 differ from comparable cases in sample 2 on important dimensions, findings based on sample 1 cannot be unequivocably generalized to all persons in King and Pierce Counties who contracted prostatic cancer between 1980 and 1982. Analysis of sample 2, however, largely remedies this problem. This sample included all newly diagnosed prostatic cancer patients entered in the CSS during the specified period. Because the number of cases the CSS fails to register is quite small, any unknown bias is likely to be minor. Except for the interaction between low income and HMO membership and its effect on mortality (income is not recorded in the CSS), the analysis of sample 2 confirms findings based on Sample 1.

Another issue concerns the unusual nature of the Group Health Cooperative of Puget Sound. The Group Health Cooperative is one of the largest and oldest HMOs in the United States, serves as a site for important research, and enjoys national recognition. The degree to which the unique qualities of the Group Health Cooperative explain its patients' apparent survival advantage is uncertain.

The categories used by the CSS to denote stage differ from those most often employed in clinical practice and investigations of prostatic cancer. Several more detailed, multidimensional staging systems are in widespread use.[13] The validity of the CSS stage as a variable representing overall disease dissemination is confirmed by its strong relationship to mortality risk in the Cox models presented above. Nevertheless, use of a different system to represent stage might have altered the results reported here.

The data reported here are not sufficiently detailed to demonstrate that the observed differences in mortality risk arise from prostatic cancer or from differential rates of success in its treatment. Like many studies of cancer mortality,[24–26] this investigation used all-cause mortality as its principal outcome variable. As in the general population of men with prostatic cancer, in which persons of advanced age are strongly represented, much of the mortality observed here is likely to have come from other causes.[27]

Despite this study's limitations, the observations reported here remain pertinent to the discussion of quality of care in HMOs versus fee-for-service plans. The finding that HMO patients are less likely to receive surgery and more likely to receive radiation therapy is consistent with reports that HMOs favor outpatient care. The results of this study contradict the belief that HMOs undertreat their patients and that low-income individuals with established health problems obtain less favorable outcomes in HMOs than in fee-for-service plans. These findings should encourage policymakers to continue viewing the HMO as a desirable alternative to traditional fee-for-service plans, particularly for low-income persons. Research published elsewhere[28] suggests that treatment in an HMO may help rem-

edy survival disadvantages faced by low-income people with cancers other than prostatic. It appears that the essential features of a well-established HMO enable persons predisposed to seeking appropriate care to do so, and that they promote access to basic medical services. Among patients similar to those studied here—those who have a major, definable health problem that is likely to coincide with a variety of eventual comorbidities—these factors appear conducive to survival. □

Acknowledgments

Collection of data was supported in part by the National Cancer Institute under Grant 5 R18 CA 32564-02. Data analysis took place under funding from the American Cancer Society (Grants PBR-11 and PDT-393).

The authors wish to thank the following people for reviewing this manuscript or assisting in its preparation: Bradford B. Burnett, Susan J. Dirks, Charles M. Haskell, Michael C. Nevitt, Betra Nicol-Blades, Ronald K. Ross, Chris Schwartz, Barbara K. Uenaka, Edward H. Wagner, and Edward H. Yelin.

References

1. Gaus CR, Cooper BS, Hirshman CG. Contrasts in HMO and fee-for-service performance. *Soc Secur Bull.* 1976;39(5):3–14.
2. Luft HS. Assessing the evidence on HMO performance. *Milbank Q.* 1980;58:501–536.
3. Luft HS. *Health Maintenance Organizations: Dimensions of Performance.* New York, NY: John Wiley & Sons; 1981.
4. Luft HS. How do health maintenance organizations achieve their "savings"?: rhetoric and evidence. *N Engl J Med.* 1978; 298:1336–1343.
5. Manning WG, Liebowitz A, Goldberg GA, et al. A controlled trial of the effect of a prepaid group practice on the use of services. *N Engl J Med.* 1984;310:1505–1510.
6. Enthoven AC. The Rand experiment and economical health care. *N Engl J Med.* 1984;310:1528–1530.
7. Brown LD. *Politics and Health Care Organization: HMOs as Federal Policy.* Washington, DC: Brookings Institution; 1983.
8. Freidson E. *Doctoring Together.* Chicago, Ill: University of Chicago Press; 1975.
9. Ware JE, Brook RH, Rogers WH, et al. Comparison of health outcomes at a health maintenance organization with those of fee-for-service care. *Lancet.* 1986;1:1017–1022.
10. Sloss EM, Keeler EB, Brook RH, et al. Effect of a health maintenance organization on physiologic health. *Ann Intern Med.* 1987;106:130–138.
11. Greenwald, HP. HMO membership, copayment, and initiation of care for cancer: a study of working adults. *Am J Public Health.* 1987;77:461–466.
12. American Cancer Society. *Cancer Facts and Figures.* Atlanta, Ga: American Cancer Society; 1990.
13. deKernion JB. Cancer of the prostate. In: Haskell CM, ed. *Cancer Treatment.* Philadelphia, Pa: WB Saunders; 1985:352–366.
14. *Cancer Patient Survival: What Progress Has Been Made?* Washington, DC: United States General Accounting Office. 1987. USGAO publication PEMD-87-13.
15. *Managed Health Care Plans: A Directory of HMOs and PPOs Offering Health Plans in Washington State.* Seattle, Wash: Puget Sound Health Systems Agency; 1985.
16. Dayal HH, Polissar L, Dahlberg S. Race, socioeconomic status, and other prognostic factors for survival from prostatic cancer. *JNCI.* 1985;74:1001–1006.
17. Smith EM, Francis AM, Polissar L. The effect of breast self-exam practices and physician examinations on extent of disease at diagnosis. *Prev Med.* 1980;9:409–417.
18. Young JL Jr, Percy CL, Asire AJ. Cancer incidence and mortality in the United States, 1973–1977. *NCI Monogr.* 1981;57: 1–187.
19. Horm JW, Asire AJ, Young JL Jr. *SEER Program: Cancer Incidence and Mortality in the United States, 1973–1981.* Washington, DC: National Cancer Institute; 1984. US Dept of Health and Human Services publication NIH 85-1837.
20. Riles LG, Pollack ES, Young JL Jr. Cancer patient survival: surveillance, epidemiology, and end results program, 1973–79. *JNCI.* 1983;70:693–707.
21. Cox DR. Regression models and life tables (with discussion). *J Royal Stat Soc (B).* 1972;34:187–220.
22. Bergner M, Bobbitt RA, Kressel S, et al. The sickness impact profile: conceptual formulation and methodological development of a health status index. *Int J Health Services.* 1976;6:393–415.
23. Cook GB, Watson FR. A comparison by age of death rates due to prostate cancer alone. *J Urol.* 1968;100:669–671.
24. Bassett MT, Krieger N. Social class and black-white differences in breast cancer survival. *Am J Public Health.* 1986;76: 1400–1403.
25. Polissar L, Sim D, Francis A. Survival of colorectal cancer patients in relation to duration of symptoms and other prognostic factors. *Dis Colon Rectum.* 1981;24:364–369.
26. Young JL, Ries LG, Pollack ES. Cancer patient survival among ethnic groups in the United States. *JNCI.* 1984;73:341–352.
27. Byar DP. The Veterans Administration comprehensive urological research group's studies of cancer of the prostate. *Cancer.* 1973;32:1126–1130.
28. Greenwald HP. *Who Survives Cancer?* Berkeley, Calif.: University of California Press; 1992.

Car Size or Car Mass: Which Has Greater Influence on Fatality Risk?

Leonard Evans, DPhil, and Michael C. Frick, MA

ABSTRACT

Objectives. Proposed increases in corporate average fuel economy standards would probably lead to lighter cars. Well-established relationships between occupant risk and car mass predict consequent additional casualties. However, if size, not mass, is the causative factor in these relationships, then decreasing car mass need not increase risk. This study examines whether mass or size is the causative factor.

Methods. Data from the Fatal Accident Reporting System are used to explore relationships between car mass, car size (as represented by wheelbase), and driver fatality risk in two-car crashes.

Results. When cars of identical (or similar) wheelbase but different mass crash into each other, driver fatality risk depends strongly on mass; the relationship is quantitatively similar to that found in studies that ignore wheelbase. On the other hand, when cars of similar mass but different wheelbase crash into each other, the data reveal no dependence of driver fatality risk on wheelbase.

Conclusions. Mass is the dominant causative factor in relationships between driver risk and car size in two-car crashes, with size, as such, playing at most a secondary role. Reducing car mass increases occupant risk. (*Am J Public Health.* 1992;82: 1105–1112)

Introduction

Relationships between occupant injury risk in a car crash and the mass of the car have been firmly established in many studies conducted over the last few decades[1], and continue to be refined.[2,3] General relationships between risk and mass can never, by themselves, establish that mass is a causative factor, because other vehicle attributes, particularly car size, are so strongly correlated with mass. Some specific findings from crash data can help illuminate causal factors. The higher rollover rates of lighter cars[4–6] appear to be caused by size rather than mass because as cars become smaller, so does their track width (the distance between the left and right wheels), with consequent reductions in rollover resistance. Another effect due to size rather than mass is the higher risk of death and injury when light cars crash into light cars compared with when heavy cars crash into heavy cars[1,7,8]; broad dynamical behavior of identical-mass cars crashing into each other should not be much influenced by their mass. Proposed increases in corporate average fuel economy are expected to lead to reductions in the mass of the average car. Some have speculated that if size were the dominant causative factor, then reductions in mass need not necessarily lead to increased occupant risk if vehicle size can be maintained constant by, for example, the use of lighter weight materials.

The purpose of the present work is to explore further the relationship between car mass and car size and its relevance to driver fatality risk. Car size is represented by wheelbase. The following four questions will be addressed, in each case for newer cars (model year 1980 or after) and for older cars (model year 1979 or before): What is the relationship between (1) mass and wheelbase? (2) driver fatality risk and wheelbase? (3) driver fatality risk and mass for cars of the same wheelbase? and (4) driver fatality risk and wheelbase for cars of similar mass?

Data

The Fatal Accident Reporting System is a computerized data file maintained by the National Highway Traffic Safety Administration containing detailed information on all traffic crashes occurring in the United States since January 1, 1975, in which anyone was killed.[9] The present investigation uses data for two-car crashes in which at least one of the drivers was killed and the mass of each car was coded. The Fatal Accident Reporting System records for the years 1975 through 1989 contain 38 465 such crashes, with 38 268 of these involving cars with coded wheelbases. Some analyses use a 25 246 crash subset, formed by excluding any crash for which more than one wheelbase was coded for either car (to reflect different versions of the same model) and any crash for which the wheelbase of each of the involved cars was the same. The analysis was performed in terms of wheelbase, rather than arguably more appropriate size indicators such as the bumper-to-bumper length, because wheelbase is the only measure of vehicle size coded directly in Fatal Accident Reporting System data.

The authors are with the Operating Sciences Department, General Motors Research Laboratories, Warren, Mich.

Requests for reprints should be sent to Leonard Evans, DPhil, Operating Sciences Department, General Motors Research Laboratories, Warren, MI 48090-9055.

This paper was submitted to the Journal September 5, 1991, and accepted with revisions March 20, 1992.

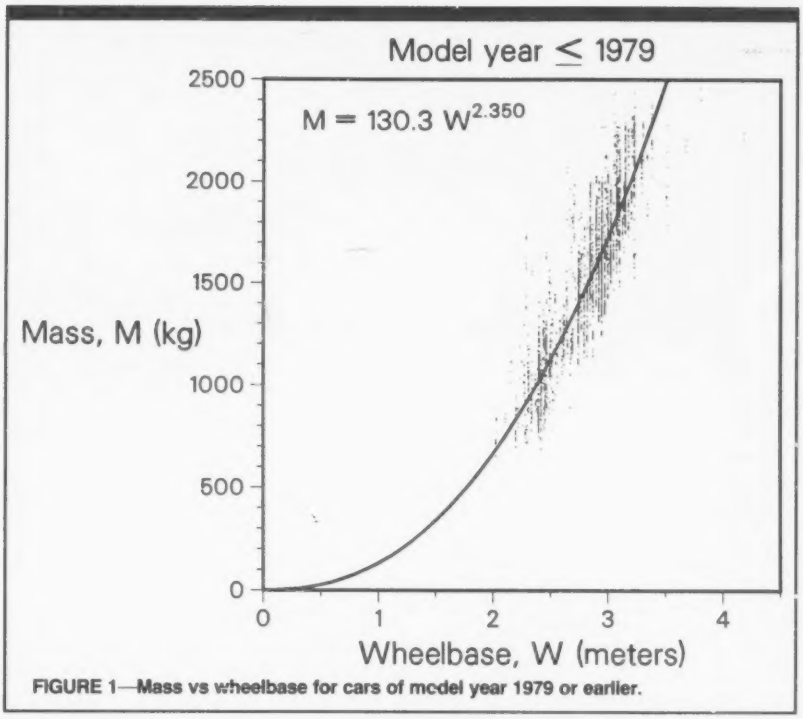


FIGURE 1—Mass vs wheelbase for cars of model year 1979 or earlier.

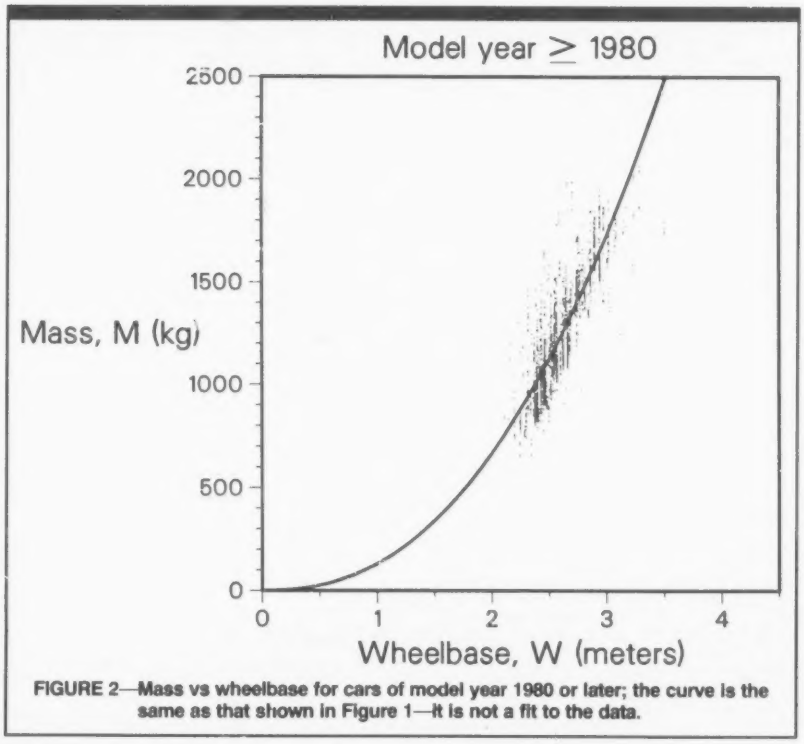


FIGURE 2—Mass vs wheelbase for cars of model year 1980 or later; the curve is the same as that shown in Figure 1—it is not a fit to the data.

Relationship between Mass and Size

To illuminate the relationship between car size and car mass, consider a hypothetical construct in which all cars are made from material with the same density. If cars were of identical shape, differing only by a scale factor, then mass would be proportional to any linear dimension to the power three. However, regardless of their size, cars must be of sufficient height to accommodate seated humans. If all cars had the same height, so that only length and breadth varied, then mass would be proportional to length (or breadth) to the power two. Real cars are likely to be intermediate between these two hypothetical cases, suggesting a relationship between mass and wheelbase of the form

$$M = \alpha W^{\beta}, \quad (1)$$

where M is the mass of the car, W is a linear dimension (other than height), and α and β are constants; the comments above lead us to expect β to be between 2 and 3.

Figure–1 shows mass vs wheelbase for cars of model year 1979 or earlier. One point is plotted for each of the 4714 unique combinations of mass and wheelbase, regardless of how many individual cars or fatalities are associated with that combination. The least squares fit of Equation 1 (in its logarithmic form) to the data gives

$$\alpha = (130.3 \pm 2.2) \quad (2)$$

and

$$\beta = (2.350 \pm 0.017). \quad (3)$$

The above value of α applies for M in kilograms and W in meters. If M is measured in pounds and W in inches, then $\alpha = 0.05124$.

The plot for cars of model year 1980 or later (3824 unique mass/wheelbase combinations) is shown in Figure 2. The least squares fit to Equation 1 gives

$$\alpha = (109.1 \pm 2.5) \quad (4)$$

and

$$\beta = (2.514 \pm 0.024). \quad (5)$$

The parameter α is smaller for the newer than for the older cars, whereas β is larger for the newer cars, in both cases the difference being statistically significant. However, because α and β are themselves so highly correlated (increasing one while decreasing the other need have only a small effect on the goodness of fit), these differences do not appear to indicate any important difference in the relationships for older and newer cars. In fact, the relationships cross at $W = 2.95$ m, $M = 1650$ kg. For $W < 2.95$ m, lower masses are calculated for newer cars (40

kg lower at $W = 2.1$ m, the minimum W in Figure 2), whereas at $W > 2.95$ higher masses are calculated for newer cars (71 kg higher at $W = 3.5$ m).

In order to examine better any changes from the older to the newer cars, the curve shown in Figure 2 is the same one shown in Figure 1 (i.e., Equation 1 with the parameters for the older cars given in Equations 2 and 3). The data for newer cars show no obvious systematic departures from the curve derived for the older cars. Thus, there does not appear to be any systematic reduction in mass apparent above the considerable scatter in the data for newer compared with older cars of the same wheelbase.

Driver Fatality Risk in Two-Car Crashes and Wheelbase

An earlier study[2] examined how the ratio of driver fatality risk in each car involved in a two-car crash depended on the ratio of the masses of the involved cars. Here we apply a parallel analysis but use the ratio of the wheelbases of the involved cars rather than the masses. In analogy with the prior mass ratio μ, we define a "wheelbase ratio," ν, for crashes between two cars of different wheelbase as

$$(6)\quad \nu = \frac{\text{Wheelbase of the Car with Longer Wheelbase}}{\text{Wheelbase of the Car with Shorter Wheelbase}}.$$

We define a driver fatality ratio, R, as

$$(7)\quad R = \frac{\text{Driver Fatalities in the Car with Shorter Wheelbase}}{\text{Driver Fatalities in the Car with Longer Wheelbase}}.$$

Older Model Year Cars

Figure 3 shows R vs ν for cars of model year 1979 or earlier. In analogy with the mass ratio study,[2] a weighted least squares fit to $\log(R) = w \log(\nu)$ is used to derive

$$(8)\quad R = \nu^w.$$

The parameter w derived from the fit to the data in Figure 3 is

$$(9)\quad w = 8.61 \pm 0.23.$$

Thus, the fatality rate increases as more than the eighth power of the wheelbase. It is rare to find so large an exponent in any relationship between observed quantities. The values of w and Δw in Equation 9, together with values derived by fitting Equation 8 to other sets of the data, are displayed in Table 1. The number of categories into which the data were divided for each analysis was the number that minimized $\Delta w/w$, subject to the restriction that no point be based on fewer than 50 fatalities.

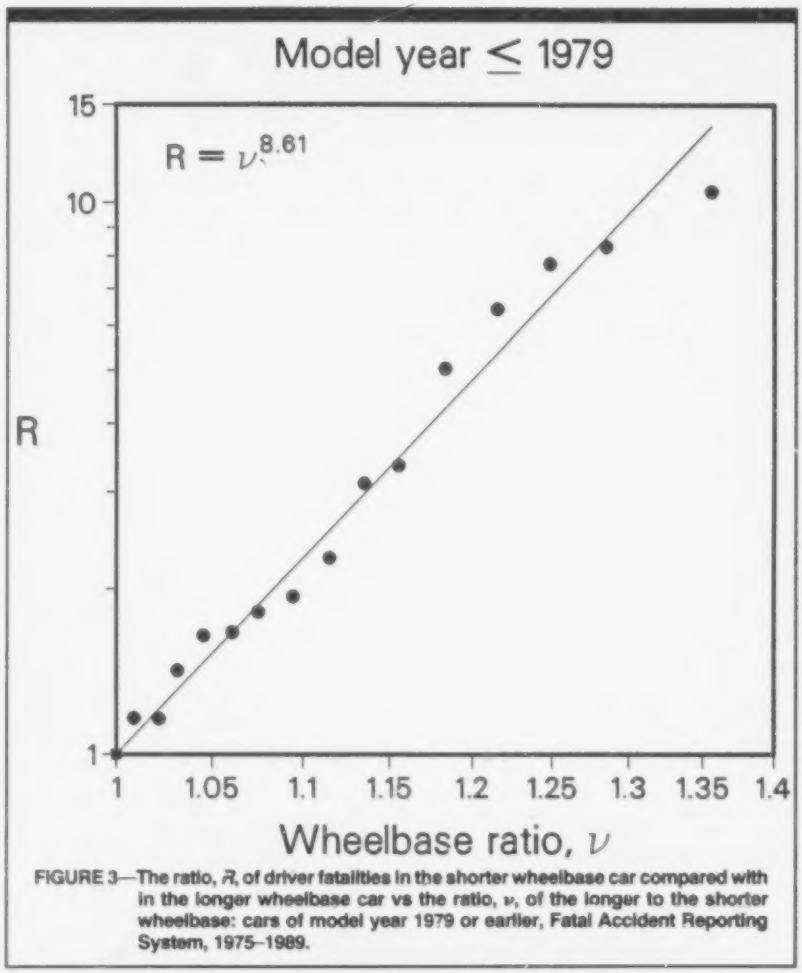


FIGURE 3—The ratio, R, of driver fatalities in the shorter wheelbase car compared with in the longer wheelbase car vs the ratio, ν, of the longer to the shorter wheelbase: cars of model year 1979 or earlier, Fatal Accident Reporting System, 1975–1989.

TABLE 1—Characteristics of the Relationship $R = \nu^w$ for Various Sets of Data

	No. of Crashes	No. of Fatalities	$w \pm \Delta w$
All cars	25 246	27 396	8.40 ± 0.16
Model year ≤ 1979 (Figure 3)	13 363	14 501	8.61 ± 0.23
Model year ≥ 1980 (Figure 4)	4351	4802	6.76 ± 0.24
Belted drivers, all cars	1397	1478	8.75 ± 0.60
Belted drivers, model year ≤ 1979	167	172	9.29 ± 0.90
Belted drivers, model year ≥ 1980	684	736	6.83 ± 0.83

Note. The numbers of categories of ν (number of points plotted in Figures 3–4) for the six cases shown are, from top to bottom, 18, 15, 12, 14, 3, and 8.

In Evans and Frick[2] the ratio, R, of driver fatality risk in the lighter of two cars compared with risk in the heavier car was related to the ratio, μ, of the mass of the heavier to that of the lighter car according to

$$(10)\quad R = \mu^w.$$

We use the same symbol R to represent

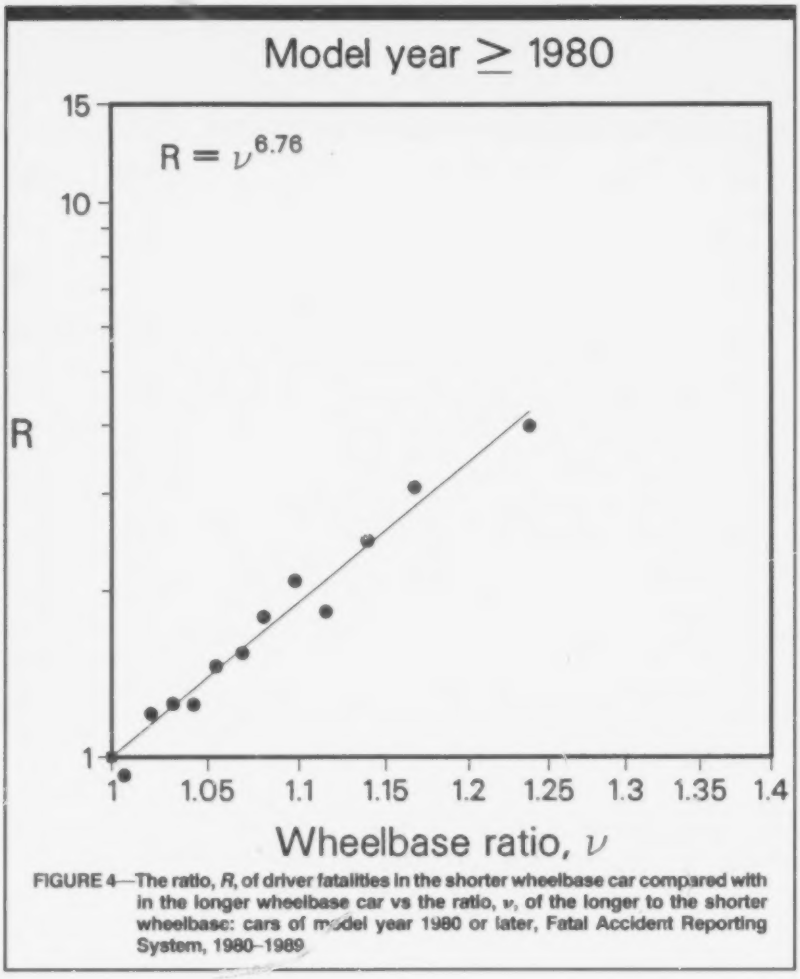


FIGURE 4—The ratio, *R*, of driver fatalities in the shorter wheelbase car compared with in the longer wheelbase car vs the ratio, *ν*, of the longer to the shorter wheelbase: cars of model year 1980 or later, Fatal Accident Reporting System, 1980–1989.

driver fatality risk ratios, although in Equation 10 the comparison is for cars of different mass, whereas in the earlier cases the comparison was for cars of different wheelbase; the context will always make it clear whether the comparison is based on mass or wheelbase.

The value of u (from Table 1 of Evans and Frick[2]) for cars of model year 1979 or earlier is

$$u = 3.73 \pm 0.04. \quad (11)$$

The much stronger relationship with wheelbase ratio than with mass ratio, as reflected by an exponent of 8.6 compared with 3.73, should not be interpreted to imply that wheelbase is a better explainer of fatality ratio than mass or that size plays a more direct causative role than weight. Rather, the relationship between fatality ratio and wheelbase ratio follows directly from the relationships between mass and wheelbase, and between fatality ratio and mass ratio, and can be inferred directly from Equation 10 (with the aid of Equations 1, 3, and 11) to give

$$R = \mu^{3.73} = \left(\frac{M_2}{M_1}\right)^{3.73} \quad (12)$$

$$= \left(\frac{\alpha W_2^{2.35}}{\alpha W_1^{2.35}}\right)^{3.73}$$

$$= (\nu^{2.35})^{3.73} = \nu^{8.8},$$

where M_2 and M_1 are the masses of the heavier and lighter cars defining the mass ratio μ and W_2 and W_1 are the corresponding wheelbases. Note how close the inferred exponent of 8.8 is to the directly obtained value of 8.6. Thus, the higher exponent for wheelbase than for mass reflects merely how steeply mass increases with increasing wheelbase. The parameter w is necessarily larger than the parameter u by about a factor of 2.35, so the difference in the magnitudes of the w and u parameters offers no illumination on the relative roles of size and weight in car safety. It seems likely that an inappropriate interpretation of larger parameters in regressions on size compared with regressions on mass could have contributed to suggestions[10] that size was a more important causative factor than mass.

Newer Model Year Cars

The data for driver fatality ratio vs wheelbase ratio for cars of model year 1980 or later are shown in Figure 4. The lower value of w, namely

$$w = 6.76 \pm 0.24, \quad (13)$$

is to be expected because the mass ratio analysis showed a lesser dependence for newer model year cars.

Applying a computation similar to that in Equation 12 to the mass ratio equation (Figure 4 of Evans and Frick[2]) and Equation 5 gives an inferred fatality risk ratio of

$$(\nu^{2.51})^{2.75} = \nu^{6.9}, \quad (14)$$

in close agreement with the directly observed value from Figure 4 of 6.8.

Effect of Safety Belt Use

The values of w and Δw in Table 1 for belted drivers indicate that the influence of wheelbase ratio on driver fatality risk is similar for belted and for all drivers (mainly unbelted). This further supports findings in three earlier studies (pp. 74 and 236 of Evans[1]; Evans and Frick[2]) that safety belt effectiveness does not depend systematically on car size. If all occupants are belted, the proportionate increase in risk in the smaller car remains similar.

How Does Fatality Risk Depend on Mass When Cars Have the Same Wheelbase?

Figures 1 and 2 show many cases of cars with the same wheelbase but different masses. This enables us to separate out effects of size and mass in a direct and straightforward manner.

Considering all two-car crashes, without regard to model year, there were 862 crashes in which the wheelbase of each of the involved cars was the same but the masses were different. In these crashes, 574 drivers in the lighter car were killed, compared with 395 driver fatalities

in the heavier car. Thus, for identical wheelbase, driver fatality risk in the lighter car was 1.45 ± 0.09 times the risk in the heavier car, where the standard error is computed using Equation 7-4 of Evans.[1] The 862 values of mass ratio are all fairly close to one, with an average value of 1.081. Substituting $\mu = 1.081$ into Equation 10 with $u = 3.53$ (Figure 1 of Evans and Frick[2]) estimates a fatality ratio of 1.32, comparable to the observed 1.45 value.

The cases in which the wheelbase of each car was identical are examined as a function of mass ratio in Figure 5. The resulting figure may be compared with many relationships between fatality ratio and mass ratio.[2] The greater noise is to be expected in view of the narrow range of values of μ; the error limits indicate one standard error for each of the plotted points. The value of u, namely 3.8 ± 0.6, is consistent with the value 3.53 derived from 38 382 crashes (Figure 1 of Evans and Frick[2]) examined as a function of mass ratio without regard to wheelbase. This comparison identifies mass unambiguously as an important factor exercising a large influence on fatality risk. Size (as represented by wheelbase) can have no effect on the relationship in Figure 5, because for each crash the wheelbase of each of the involved cars is the same. If the analysis is confined to cars of model year 1979 or earlier, 573 of the original 862 crashes are preserved. These give $u = 4.3 + 0.6$, which is consistent with the value 3.73 for 23 012 crashes without regard to wheelbase.[2]

Confining the analysis to crashes between two cars of model year 1980 or later leaves 168 crashes that led to 102 driver fatalities in lighter cars compared with 88 in heavier cars, for a ratio of 1.16 ± 0.17. Substituting $\mu = 1.051$, the average mass ratio, into Equation 10 with $u = 2.75$ (Figure 4 of Evans and Frick[2]) computes an expected driver fatality ratio of 1.15. The difference between the computed ratio and unity is too small to be detected reliably with the available sample size (the closeness of the observed and expected values is fortuitous given the large uncertainty).

In order to increase sample sizes, the restriction that the wheelbases be equal was replaced by a restriction that they not differ by more than some small amount. As this amount is increased, sample sizes increase, but so also does the possibility that the difference in wheelbase is itself influencing the driver fatality risk ratio. There were 280 crashes involving two cars

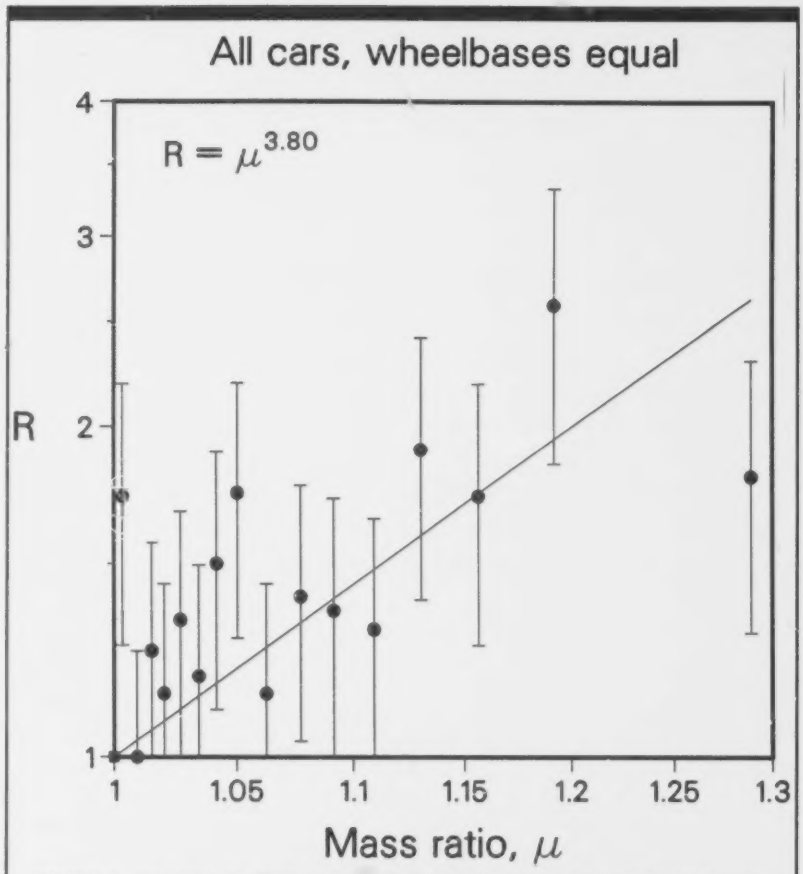


FIGURE 5—The ratio, *R*, of driver fatalities in the lighter compared with in the heavier car vs the ratio, μ, of the mass of the heavier to the mass of the lighter car for crashes between cars with identical wheelbase: data for cars of any model year, Fatal Accident Reporting System, 1975–1989.

of model year 1980 or later whose wheelbases differed by less than 1 cm, leading to 183 driver fatalities in the lighter cars compared with 133 in heavier cars. Very little of this difference seems attributable to the differences in the wheelbases, because a 1-cm difference in a 2-m-wheelbase small car gives $\nu = 1.005$, which (from the relationship in Figure 4) would produce a 3.4% effect on driver fatality ratio. This is small compared with the (38 ± 16)% difference observed.

Figure 6 shows fatality risk ratio vs mass ratio for the 280 crashes in which the wheelbases of the two cars involved differed by less than 1 cm. The value of u, 3.76 ± 0.89, is consistent with the value of 2.75 ± 0.11 for cars of model year 1980 or later.[2] Given the uncertainty, the most that can be concluded is that when the wheelbase remains fairly constant, the data provide no indication that fatality risk ratio increases with mass ratio less strongly than when wheelbase is ignored. If wheelbase and mass were exercising a shared influence, then when wheelbase is fixed one would expect the influence of mass to be reduced.

The analyses for cars of all model years, for older cars, and for newer cars all consistently lead to the same conclusion. When wheelbase is held constant, mass influences driver fatality ratio just as strongly as when wheelbase varies. This provides strong evidence that mass, as such, has a large influence on driver fatality risk.

How Does Fatality Risk Depend on Wheelbase When Cars Have Similar Mass?

The above analysis suggests a corresponding one to address the question "If the mass is the same, how does wheelbase ratio affect fatality risk?" While there are usable sample sizes of cars coded with identical wheelbase but different mass, there were only eight crashes between

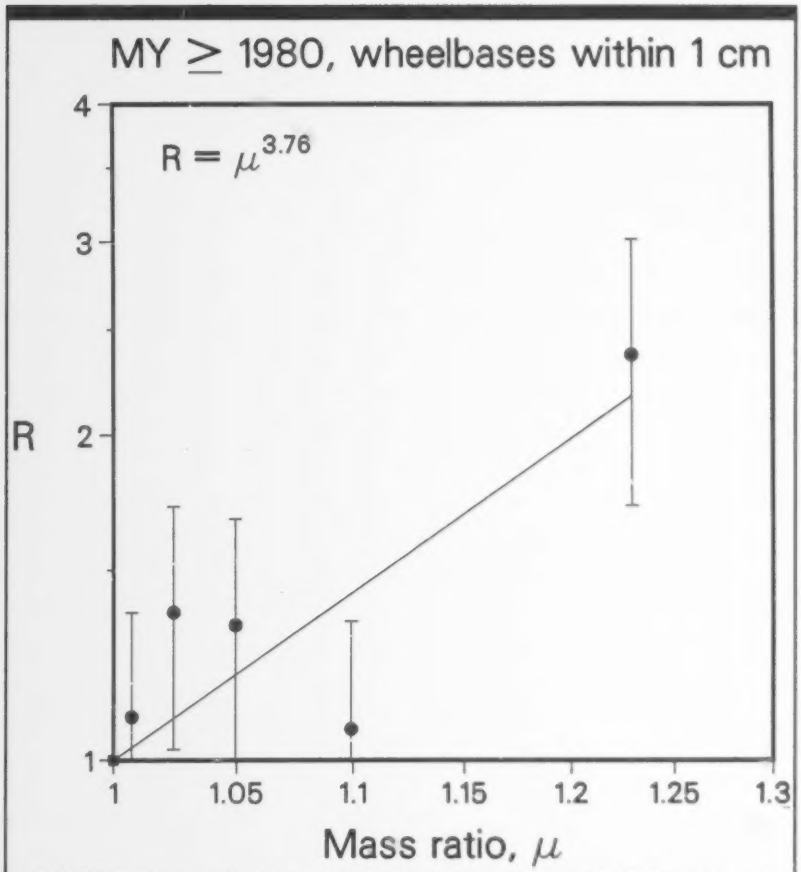


FIGURE 6—The ratio, *R*, of driver fatalities in the lighter compared with in the heavier car vs the ratio, μ, of the mass of the heavier to the mass of the lighter car for crashes between cars with wheelbases differing by less than 1 cm: cars of model year 1980 or later, Fatal Accident Reporting System, 1980–1989.

TABLE 2—Crashes between Cars of Similar Mass: Examination of the Influence of Wheelbase for Cars of Model Year 1979 or Earlier

Maximum Difference between Car Masses, kg	No. of Crashes	Fatalities in: Shorter Car	Fatalities in: Longer Car	Ratio ± SE	$\bar{\nu}$
20	321	173	185	0.935 ± 0.099	1.044
30	470	257	275	0.935 ± 0.081	1.043
40	633	351	365	0.962 ± 0.072	1.043
50	780	438	455	0.963 ± 0.064	1.043
70	1084	610	622	0.981 ± 0.056	1.043
100	1549	876	878	0.998 ± 0.048	1.044
150	2337	1339	1302	1.028 ± 0.040	1.046
200	3175	1850	1733	1.068 ± 0.036	1.050
250	4022	2416	2109	1.146 ± 0.034	1.054

cars of the same mass but different wheelbase. As the criterion of equal mass is too restrictive, we consider immediately crashes between cars with masses the same to within some small amount and examine driver fatalities in the shorter wheelbase car compared with those in the longer wheelbase car.

The results for cars of model year 1979 or earlier (Table 2) show no hint of any dependence of fatality risk on wheelbase when the masses of the cars are similar. When the masses are within 20 kg of each other, the average wheelbase ratio is 1.044. If the relationship in Figure 3 reflected a pure influence of wheelbase ratio (rather than a surrogate for mass) on fatality risk, then substituting $\nu = 1.044$ should generate the observed driver fatality risk ratio. The computed value of 1.45 should be contrasted to the observed value (not different from 1). Recall that applying the corresponding calculation when the wheelbase was fixed led to reasonable agreement between predicted and observed fatality ratios. It is only when the masses differ by more than 150 kg that the fatality ratio begins to depart clearly from unity. This difference is interpreted as being due to the strong correlation between wheelbase and mass.

For cars of model year 1980 or later (Table 3), fatalities in the shorter car exceed those in the longer car for all mass differences tabulated. However, for no case when the masses are within 150 kg of each other does the difference approach any normally quoted level of statistical significance; indeed, as the sample sizes increase, the ratio of fatalities in the shorter car to those in the longer approaches closer to one until the effect of mass begins to influence strongly results at mass differences in excess of 200 kg. Substituting 1.043 into the relationship in Figure 4 computes a ratio of 1.33. None of the observed ratios comes close to such a value, again demonstrating that the relationship in Figure 4 is not showing the influence of size, but of mass.

The results in Tables 2 and 3 consistently show that when cars of similar mass but different wheelbase crash into each other, the wheelbase does not produce any clearly observable effect on driver fatality risk ratio.

Discussion

Examining crashes between cars with the same wheelbase provides definitive empirical evidence that mass is the main causative factor in the strong relationship between driver fatality risk and car mass in two-car crashes. The relationship between driver fatality risk and mass ratio when cars of equal or similar wheelbase crash into each other is quantitatively similar to the relationship with mass ratio when cars in general crash into each other, regardless of wheelbase. On the other hand, when cars of similar mass crash into each other, no effects dependent on wheelbase are found.

There are compelling physical reasons why increased car mass should reduce occupant risk in essentially all types of crashes. In the case of two-car crashes, the effect flows directly from Newton's

laws of motion, which require that when two objects collide, the speed change each undergoes is in inverse proportion to its mass. It is this speed change that generates the forces that lead to death and injury. Because the speed change decreases with increased mass by amounts unrelated to size or shape, increased car mass intrinsically provides increased protection to the occupants of a car when that car strikes another car, or any other vehicle, or any other object.

There is evidence that car size does play a role in two-car crashes. The strongest is that when cars of the same mass crash into each other, fatality and injury risk is lower for two heavier cars crashing into each other than for two lighter cars crashing into each other.[1,3,7,8] As the broad dynamical characteristics of a crash between two identical cars should not depend on their mass, the explanation of this finding would appear to be in terms of car size rather than mass, even though the relationships are reported in terms of mass. If one accepts the strongest relationship published (doubling the mass halves the risk), and the relationship between car mass and wheelbase derived here, then a 4% larger wheelbase would generate about a 10% lower driver fatality risk. This is just below the threshold of detectability, although such a difference is arguably a more harmonious interpretation of Table 3 than the equally defensible interpretation of no effect.

All the analyses in this paper have related fatality risk ratios to ratios of wheelbase or mass. These analyses in themselves cannot exclude the possibility that larger size in one involved car could generate equal proportionate risk reductions for both involved drivers, with no consequent effect on the fatality risk ratio. Based on the finding that the effect of mass is similar whether wheelbase is fixed or allowed to vary, and the much lesser magnitude of other effects that can be attributed to size, we consider such effects to be, at most, small compared with the large and unmistakable effect of mass. We consider that car size analyses that focus on car mass,[2,3] are focusing on the single variable that best captures most, but certainly not all, of the safety changes that occur as a function of changing vehicle size.

The conclusion that size has no more than a small influence on fatality risk in two-car crashes suggests that size may have a correspondingly small effect on injury outcome (as distinct from crash involvement) in single-car crashes. If so, then a substantial portion of relationships between driver fatality risk and car mass in single-car crashes[1,11] may likewise be due to mass, as such. Conceptually, the role of mass in single-car crashes has often been misunderstood because barrier crash tests so often come to mind as representing single-car crashes. A barrier crash test is so inadequate a representation of real-world crashes that it is not surprising that no relationship between barrier crash test performance and real-world performance has ever been reliably observed. Unlike the nominal design features of the barrier, the average object struck in the real world is not infinitely stiff, infinitely massive, or infinitely unbreakable.

In any large set of crashes, many of the objects struck by vehicles will move, dent, deform, or tear. Each of these effects reduces forces on occupants. The heavier the striking vehicle, other factors being equal, the more likely is a struck object (such as a tree, the most common fixed object in fatal crashes) to break. If the vehicle breaks or uproots the tree, rather than bouncing back from it, risk to occupants is dramatically reduced. Even the proverbial brick wall from which a light car will bounce back may still be penetrated by a heavier vehicle, with beneficial effects to its occupants. Observed relationships between car-driver risk in car–truck crashes[2,11] may arise in part because increased car mass will, on average, produce greater damage to the impacted area of the truck, with consequent reductions in the forces on the car driver.

Risks in single-car crashes, two-car crashes, and crashes between cars and other vehicles all increase with decreasing car mass. Many of the mechanisms flow from the most basic physical laws involving mass in dynamical equations, whereas others, such as substantially higher rollover crash risk in lighter cars, flow from the close relationship between car mass and car size. All effects combine to generate a strong relationship between overall risk and car mass. It is not surprising that all of the many studies performed using credible methodologies have found large relationships between overall risk and car mass. A recent claim[12] that mass is unimportant seems just one more example of "the triumph of zeal over science, or perhaps even common sense"[1(p83)]; a few regression parameters backed up by no raw data are presented to support this claim, which implies that all prior work is incorrect and that cars do not obey the laws of physics after all.

TABLE 3—Crashes between Cars of Similar Mass: Examination of the Influence of Wheelbase for Cars of Model Year 1980 or Later

Maximum Difference between Car Masses, kg	No. of Crashes	Fatalities in: Shorter Car	Fatalities in: Longer Car	Ratio ± SE	$\bar{v}$
20	166	98	87	1.126 ± 0.166	1.043
30	246	144	128	1.125 ± 0.137	1.042
40	326	183	175	1.103 ± 0.115	1.042
50	405	239	212	1.127 ± 0.106	1.041
70	575	327	310	1.055 ± 0.084	1.042
100	831	470	456	1.031 ± 0.068	1.042
150	1236	706	672	1.051 ± 0.057	1.045
200	1642	949	876	1.083 ± 0.051	1.048
250	2053	1223	1054	1.160 ± 0.049	1.051

Conclusions

As a result of this study, we made the following conclusions: First, when cars of similar wheelbase but different mass crash into each other, driver fatality risk is higher in the lighter car. The large difference agrees quantitatively with predictions of relations derived only in terms of mass, without considering wheelbase. Second, when cars of similar mass but different wheelbase crash into each other, driver fatality risk in the shorter and longer wheelbase cars are similar. Finally, mass is the dominant causative factor in the large dependence of driver fatality risk on mass in two-car crashes, with size playing at most a secondary role. □

References

1. Evans L. *Traffic Safety and the Driver.* New York, NY: Van Nostrand Reinhold; 1991.
2. Evans L, Frick MC. Mass ratio and relative driver fatality risk in two-vehicle crashes. *Accident Analysis Prev.* In press.
3. Evans L, Frick MC. Driver fatality risk in two-car crashes—dependence on masses of driven and striking car. In: *Proceedings of the Thirteenth International Technical Conference on Experimental Safety Vehicles, Paris, France, 4–7 November 1991.*

Washington, DC: National Highway Traffic Safety Administration. In press.
4. Partyka SC, Boehly WA. Passenger car weight and injury severity in single vehicle nonrollover crashes. In: *Proceedings of the Twelfth International Technical Conference on Experimental Safety Vehicles, Gothenburg, Sweden, 29 May–1 June 1989.* Washington, DC: National Highway Traffic Safety Administration; 1990:507–519.
5. Mengert P, Salvatore S, DiSario R, Walter R. *Statistical Estimation of Rollover Risk.* Cambridge, Mass: National Highway Traffic Safety Administration, Transportation Systems Center; 1989. Report DOT-HS-807-446/DOT-TSC-NHTSA-89-3.
6. Kahane CJ. *An Evaluation of Door Locks and Roof Crush Resistance of Passenger Cars—Federal Motor Vehicle Safety Standards 206 and 216.* Washington, DC: National Highway Traffic Safety Administration; 1989. Report DOT HS 807-489.
7. Evans L, Wasielewski PF. Serious or fatal driver injury rate versus car mass in head-on crashes between cars of similar mass. *Accident Analysis Prev.* 1987;19: 119–131.
8. Klein TM, Hertz E, Borener S. *A Collection of Recent Analyses of Vehicle Weight and Safety.* Washington, DC: National Highway Traffic Safety Administration; 1991. Report DOT HS 807-767.
9. National Highway Traffic Safety Administration. *Fatal Accident Reporting System 1989.* Washington, DC: National Highway Traffic Safety Administration; 1991. Document DOT HS 807-693.
10. O'Neill B, Joksch H, Haddon W Jr. Empirical relationships between car size, car weight and crash injuries in car-to-car crashes. In: *Proceedings of the Fifth International Technical Conference on Experimental Safety Vehicles, London, UK, 4–7 June 1974.* Washington, DC: National Highway Traffic Safety Administration; 1975:362–368.
11. Evans L. Driver fatalities versus car mass using a new exposure approach. *Accident Analysis Prev.* 1984;16:19–36.
12. Robertson LS. How to save fuel and reduce injuries in automobiles. *J Trauma.* 1991;31:107–109.

Call for Abstracts for Epidemiology Late-Breaker Sessions

Exchange Session

The Epidemiology Section will sponsor a "late-breaker" Epidemiology exchange on Wednesday, November 11, 1992, during APHA's 1992 annual meeting in Washington, DC. The exchange will provide a forum for presentation of investigations or methods that have been conceived, conducted, and/or concluded so recently that members could not meet the deadline for abstract submission to other Epidemiology sessions. Abstracts should report on work conducted during the last 6–12 months.

Abstracts of fewer than 200 words (any format) and a stamped, self-addressed return envelope should be submitted to Polly A. Marchbanks, PhD, Chief, EIS Program, Epidemiology Program Office, Mailstop C08, Centers for Disease Control, Atlanta, GA 30333; (404) 639-3588.

Abstracts must be *received* by September 28, 1992. Decisions will be made by October 5, 1992.

Poster Session

The Epidemiology Section will again be sponsoring a late-breaker poster session at the annual meeting. The session will take place on Wednesday, November 11, 1992. Work completed in 1992 is eligible for consideration. Submit an abstract of fewer than 200 words (any format) and a *return envelope* to Cathey Falvo, MD, MPH, Graduate School of Health Science, Munger Pavilion, New York Medical College, Valhalla, NY 10595; tel. (914) 993-4323.

Abstracts must be *received* by October 2, 1992. Decisions will be mailed October 5, 1992. Students and recent graduates are particularly encouraged to submit abstracts.

The Development of Registries for Surveillance of Adult Lead Exposure, 1981 to 1992

Michael E. Baser, PhD

ABSTRACT

Objectives. Since 1981, 15 states have established registries for surveillance of adult lead absorption, primarily based on reports of elevated blood lead levels from clinical laboratories. I review the status of the registries and recommend steps for further development.

Methods. Companies reported to the New York registry are compared with those cited by the Occupational Safety and Health Administration (OSHA). I present data on US workers and plants with potential lead exposures and blood tests, as well as review registries' reporting requirements.

Results. Registries identify many companies not cited by the Occupational Safety and Health Administration, but underreporting occurs because (1) reporting is usually not required from laboratories outside the state, (2) most registries use a blood lead reporting level of 1.21 μmol/L, which excludes many exposed workers, and (3) many companies with potential exposures do not have routine monitoring programs.

Conclusions. Registries' reporting requirements and procedures should be standardized, including a blood lead reporting level of 0.72 μmol/L. Elevated blood lead levels should be a reportable condition nationwide, and a comprehensive national surveillance system should be established: clinical laboratories should be required to report cases to those states with lead registries or directly to the national adult lead registry. (*Am J Public Health.* 1992;82:1113–1118)

Introduction

Lead adversely affects the hematologic, central and peripheral nervous, cardiovascular, renal, and reproductive systems.[1] A variety of surveillance methods have been proposed to identify, first, morbidity and mortality potentially associated with lead exposures and, second, companies or workers with lead exposures.[2–7] The National Occupational Exposure Survey (NOES),[8] conducted from 1981 to 1983, estimated that almost 1.4 million workers were exposed to inorganic lead or lead compounds (Table 1). Many industries have airborne lead concentrations greater than the permissible exposure limit (PEL) set by the Occupational Safety and Health Administration (OSHA), which is 50 $\mu g/m^3$ averaged over an 8-hour work shift.[9,10] Furthermore, the prevalence and severity of exposures in a number of industries with high lead exposures have not declined in recent years.[10]

The OSHA lead standard requires that companies provide industrial hygiene and medical surveillance programs (including measurement of blood lead levels) when airborne lead concentrations exceed 30 $\mu g/m^3$ averaged over an 8-hour work shift.[9] A worker must be removed from exposure if (1) the mean blood lead level of the worker's last three tests, or the mean of all of the worker's tests over the previous 6 months, is greater than or equal to 2.42 μmol/L (1 μmol/L = 20.72 μg/dL); or (2) any test is greater than or equal to 2.90 μmol/L; or (3) the worker's medical condition contraindicates continued lead exposure.

Methods and Results

In 1981, the New York State Health Department began population surveillance of adult lead exposure, using blood lead results from industrial medical surveillance programs and other sources.[7] State statutes require clinical laboratories, employers, and health facilities to report individuals with elevated levels of heavy metals in blood and urine samples. State agencies provide follow-up, which includes education for employers, workers, and clinicians, industrial hygiene evaluations, and medical consultation.

Adult lead registries have been established in 15 states, primarily based on reporting from clinical laboratories (Table 2). Through 1990, the New York registry had received reports on 4735 workers from 695 companies or work sites; the New Jersey registry, 2851 workers from 271 companies; the California registry, 2779 workers from 250 companies;[11] and the Texas registry, 1116 workers from 101 companies. Many of these companies had employees with markedly elevated blood lead levels. Of the 695 work sites reported to the New York registry, 197 (28.3%) had at least one employee with blood lead levels exceeding the OSHA medical removal protection levels. These included 15 of 34 heavy construction companies (44.1%; Standard Industrial Classification[12] [SIC] 16), 27 of 61 special trade contractors (44.3%; SIC 17), 22 of 43 primary metal industries (51.2%; SIC 33), 17 of 32 fabricated metal product companies (53.1%; SIC 34), 13 of 25 scrap

This article was partially written while the author was with the Department of Environmental Health Sciences, The Johns Hopkins School of Hygiene and Public Health, Baltimore, Md.

Requests for reprints should be sent to Michael E. Baser, PhD, Neurogenetics Laboratory, Room B-216, Cedars-Sinai Medical Center, Los Angeles, CA 90048.

This paper was submitted to the Journal June 13, 1991, and accepted with revisions December 31, 1991.

TABLE 1—Workers and Plants with Potential Inorganic Lead Exposures,[a] by Major Industrial Group, from the National Occupational Exposure Survey (NOES), 1981 to 1983

NOES Major Industrial Group (Standard Industrial Classification)	Workers					Plants				
	Total	With Potential Lead Exposure[b]		With Blood Tests[c]		Total	With Potential Lead Exposure[b]		With Blood Tests[c]	
	No. (SE)	No. (SE)[d]	% of Total Workers	No. (SE)[d]	% of Workers with Lead Exposure	No. (SE)[d]	No. (SE)[d]	% of Total Plants	No. (SE)[d]	% of Plants with Lead Exposure
Agricultural services (7)	110 692 (29 333)	0 ...	0.0	0 ...	...	5633 (1557)	0 ...	0.0	0 ...	...
Oil and gas extraction (13)	414 284 (101 840)	45 600 (6840)	11.0	0 ...	0.0	9662 (2132)	2064 (722)	21.4	0 ...	0.0
Construction (15–17)	3 072 049 (132 323)	198 790 (19 879)	6.5	1812 (652)	0.9	98 791 (2409)	12 079 (2657)	12.2	183 (121)	1.5
Manufacturing (20–39)	19 261 829 (551 881)	900 643 (63 045)	4.7	318 147 (28 633)	35.3	191 266 (3668)	39 716 (8355)	20.8	5953 (1548)	15.0
Transportation (40–49)	3 160 926 (293 218)	113 464 (13 616)	3.6	32 085 (5454)	28.3	59 465 (3540)	10 935 (2406)	18.4	873 (384)	8.0
Wholesale/retail trade (50–59)	1 533 201 (142 606)	15 103 (3021)	1.0	3245 (1006)	21.5	61 051 (3970)	3643 (1093)	6.0	537 (268)	14.7
Services (70–79)	2 197 247 (133 603)	79 573 (10 344)	3.6	13 556 (2847)	17.0	75 835 (3590)	8130 (1951)	10.7	319 (185)	3.9
Health services (80)	3 658 805 (351 918)	24 612 (4430)	0.7	21 994 (3959)	89.4	7067 (851)	1514 (575)	21.4	1347 (525)	89.0
Total	33 409 033 (762 445)	1 377 785 (96 445)	4.1	390 839 (31 267)	28.4	508 770 (8254)	78 081 (10 150)	15.3	9212 (2119)	11.8

Sources. Data from NOES[9] and Standard Industrial Classification Manual.[12]

[a]Inorganic lead in lead ore, lead compounds/complexes, lead alloys, and lead in welding, brazing, and soldering. An estimated 222 207 ± 22 221 workers were exposed to organolead compounds.

[b]Potential or inferred exposure is annual average exposure ≥ 30 minutes/week, or usage ≥ once/week for 90% of weeks/year.

[c]"Blood tests" is an affirmative answer to the question of blood tests provided to selected or all employees, with the exception of tests provided only to executive or management employees. The question does not specifically refer to lead.

[d] ... = not available. SEs were estimated using coefficients based on a nonlinear regression of 792 SEs for different-sized estimates from 44 SIC groups; at present, SEs cannot be estimated for zero values or for percentages.

and waste material companies (52.0%; SIC 5093), and 20 of 84 automotive repair shops (23.8%; SIC 7539).

Although many workers and companies are reported to registries, registries underestimate the prevalence of occupational lead exposure because only a small fraction of companies with potential lead exposures perform routine biological monitoring. A survey of California lead-using companies estimated that only 2.6% ever performed environmental monitoring and only 1.4% had routine biological monitoring.[13] Using different methods in a national sample,[8] the NOES estimated that 11.8% of companies with potential inorganic lead exposures had blood tests performed on employees (Table 1). The prevalence of routine biological monitoring is especially low in small companies and the construction industry,[8,13,14] which has a permissible exposure limit of 200 $\mu g/m^3$ and no OSHA requirements for routine air or blood sampling.[15] The NOES estimated that only 0.1% of 198 790 lead-exposed construction workers received blood tests and that there were no blood tests for exposed workers in oil and gas extraction (SIC 13), auto repair, services, and garages (SIC 75), and miscellaneous repair services (SIC 76).

Lead registries complement OSHA inspections in two ways. First, registries detect many companies not cited by OSHA,[7,16] whose inspection program focuses on larger manufacturing companies with well-recognized lead hazards. Through 1990, only 76 of the 695 work sites (10.9%) reported to the New York registry were also cited by OSHA. These included 62 of 187 manufacturing companies (33.2%) but only 14 of 508 nonmanufacturing companies (2.8%) ($P < .001$, chi-square test). Second, registry data and OSHA inspection results can be used to target interventions because companies detected by both methods have particularly high exposures.[12] Through 1990, 580 New York companies were cited by OSHA for violations of the federal lead standard. Of 76 cited companies that were also reported to the New York registry, 49 (64.5%) were cited for airborne lead concentrations above the permissible exposure limit, compared with 70 of 504 companies (13.9%) cited by OSHA but not reported to the registry ($P < .001$, chi-square test).

Recommendations for Registry Development

The reporting requirements and procedures vary among states (Table 2) and should be standardized to ensure wider and more uniform reporting and a common basis for interventions and data comparisons.

1. *Blood lead reporting level.* A level of 0.72 μmol/L or lower should be adopted, as four states have done. The probability of detecting excessive expo-

TABLE 2—Reporting Requirements for Adult Lead Registries, 1992

State (Start Date)	Contact Person, Address, and Telephone Number	Blood Lead Reporting Level (μmol/L)	Other Heavy Metals	Required to Report to Registry, Days	Age at Which Person Is Defined as an Adult, y	Reporting Also Required from Individual Clinicians and Health Facilities?
Alabama (11/90)	Charles Woemle, MD, MPH State Epidemiologist Department of Public Health 434 Monroe Street Montgomery, AL 36130 (205) 242-5131	0.72	None	7	18	Yes
California (1/87)	Ana Maria Osorio, MD, MPH Occupational Health Program California Department of Health Services 2151 Berkeley Way, Room 504 Berkeley, CA 94704 (510) 540-2115	1.21	None	3	16	No
Colorado (10/88)	Jane McCammon Epidemiology Division Colorado Department of Health 4210 E 11th Avenue Denver, CO 80220 (303) 331-8330	1.21	Mercury Arsenic Cadmium	30	18	No
Connecticut (1/91)[a]	Narda Tolentino, MPH Environmental Epidemiology and Occupational Health Connecticut Department of Health Services 150 Washington Street Hartford, CT 06106 (203) 566-8167	1.21	None	Not stated	16	Yes
Illinois (4/90)	Melinda Lehnherr Occupational Disease Registry Division of Epidemiologic Studies Illinois Department of Public Health 605 W Jefferson Springfield, IL 62761 (217) 785-1873	1.21	None	7	16	No
Iowa[b] (10/90)	Russell Currier, DVM, MPH Environmental Epidemiology Section Iowa Department of Public Health Lucas State Office Building Des Moines, IA 50319 (515) 281-5643	All levels reportable	Mercury Arsenic Cadmium	Not stated	15	Yes
Maryland (2/88)	Ellen Coe, RN, MPH Health Registries Division Maryland Department of the Environment 2500 Broening Highway Baltimore, MD 21224 (410) 631-3851	1.21	Mercury Arsenic Cadmium	5	18	No
Massachusetts (4/91)	Richard Rabin Division of Occupational Hygiene Massachusetts Department of Labor and Industries 1001 Watertown Street West Newton, MA 02165 (617) 969-7177	0.72	None	7	15	No
Michigan (3/92)	Flint Watt, MBA, CIH Division of Occupational Health Michigan Department of Public Health 3423 N Logan Boulevard Lansing, MI 48909 (517) 335-8250	1.93	None	3	16	No

(Continued on next page)

TABLE 2—(Continued)

State (Start Date)	Contact Person, Address, and Telephone Number	Blood Lead Reporting Level (μmol/L)	Other Heavy Metals	Required to Report to Registry, Days	Age at Which Person Is Defined as an Adult, y	Reporting also Required from Individual Clinicians and Health Facilities?
New Jersey (10/85)	Barbara Gerwel, MD Occupational Disease Prevention Program New Jersey Department of Health C N 360, John Fitch Plaza Trenton, NJ 08625 (609) 984-1863	1.21	Mercury Arsenic Cadmium	Immediately	16	Yes
New York (9/81)	Robert Stone, PhD Division of Occupational Health New York State Department of Health 2 University Place, Room 155 Albany, NY 12203-3313 (518) 458-6228	1.21[d]	Mercury Arsenic Cadmium	10	18	Yes
Oregon (9/90)	Jane Gordon, PhD Deputy State Epidemiologist Oregon Health Division 800 NE Oregon Street, #21 Portland, OR 97232 (503) 731-4025	1.21	None	7	18	No
Texas (10/85)	Teresa Wills Texas Department of Health 1100 West 49th Street Austin, TX 78756 (512) 458-7268	1.93	None	Not stated	15	Yes
Utah[b] (1/90)	Denise Beaudoin, MD, MSPH Bureau of Epidemiology Utah Department of Health P.O. Box 16660 Salt Lake City, UT 84116-0660 (801) 538-6191	0.72[e]	None	60	18	Yes
Wisconsin (12/87)	Larry Hanrahan, MS Bureau of Epidemiology Wisconsin Department of Health and Social Services One W Wilson Street, Box 309 Madison, WI 53701 (608) 267-7173	1.21	None	2	16	Yes
CDC-NIOSH (1/87; data from existing registries)	Paul J. Seligman, MD, MPH Surveillance Branch National Institute for Occupational Safety and Health 4676 Columbia Parkway, R-21 Cincinnati, OH 45226 (513) 841-4353	...	None	...	...	...

Sources. State registries and the Centers for Disease Control's National Institute for Occupational Safety and Health (CDC–NIOSH) registry.
[a]Pilot data collected from the state public health laboratory, 1989–1990.
[b]Occupational follow-up planned or under consideration.
[c]Lowered from 1.21 μmol/L, 4/92.
[d]Lowered from 1.93 μmol/L, 10/86 (1 μmol/L = 20.72 μg/dL).
[e]Lowered from 1.45 μmol/L, 12/91.

sures is related to the frequency of exposure measurement.[17] Because many companies with lead exposures do not have biological monitoring programs, a relatively low blood lead level that is incidentally reported to a registry can be indicative of higher exposures.

Most states currently have a blood lead reporting level of 1.21 μmol/L, and the number of reported cases will increase if the reporting level is set at 0.72 μmol/L. For example, 49% of the workers reported to the Alabama registry in 1991 had initial blood lead levels from 0.72 to 1.20 μmol/L. Additional personnel should be provided to ensure adequate follow-up.

2. *Other information on lead exposures.* To augment reporting based on blood lead levels, registries should reference information on lead-using companies, OSHA inspection results, and workers' compensation claims.[7,13,16,18] Some registries estimate the prevalence of elevated lead absorption by determining the

number of potentially exposed workers, using a questionnaire; the California questionnaire is a good model. The Iowa registry requires the reporting of the results of all blood tests. Work histories should be elicited from workers aged 15 and older, in order to include those who leave school before age 18 and students with summer or part-time jobs.

3. *Other heavy metals.* Increased absorption of mercury, arsenic, and cadmium should be reported. Through 1990, the New York registry received reports on 1505 mercury cases from 70 companies (including many workers from a chlor-alkali plant), 151 arsenic cases from 22 companies, and 63 cadmium cases from 18 companies.

4. *Response time.* To expedite follow-up in the most serious cases, blood lead levels above 2.42 μmol/L (the OSHA biological limit value) should be reported within 24 hours. Through 1990, only 24.7% of New York tests above 2.42 μmol/L were reported within 10 days and 63.7% within 30 days; only 19.1% of all tests were reported within 10 days, the statutory requirement. In 1987, only 9% of California tests were reported within 3 days.[19]

5. *Reporting sources.* Clinicians are required by many states to report lead or heavy metals poisoning[20] and should also report directly to registries. In practice, employers, health facilities, and clinicians send samples to laboratories, which then send the results to registries. In New York, through 1990 93.4% of workers and 95.9% of companies have been reported by laboratories, alone or jointly with other sources. Laboratories should be audited to ensure that all cases have been reported, as the New York registry has done.

6. *Location of reporting sources.* Reporting should be sought from out-of-state laboratories because clinical laboratories do not have strictly local clients. Of the 15 state registries, only New York requires reporting from all laboratories that analyze blood from state residents. Thirty-three of the 72 laboratories reporting to the New York registry are located outside the state; through 1990, 80.5% of workers and 77.0% of companies have been reported from these laboratories. An estimated 47% of blood lead samples from biological monitoring programs in California are analyzed by laboratories outside the state.[13]

Several methods can be used to broaden reporting from clinical laboratories. All states should designate elevated blood lead levels as a reportable condition and, as in New York, should use or create licensing authority to require reporting from both in-state and out-of-state laboratories. Alternatively, federal legislation could be enacted for these purposes. Blood lead levels from residents of states without registries should be reported directly to the national adult lead registry, which could then share the data with state health and labor departments. The national registry, at the Centers for Disease Control's National Institute for Occupational Safety and Health (CDC-NIOSH), uses data from existing registries (Table 2).

7. *Identification of laboratories.* Commercial, public health, and university hospital laboratories that analyze lead in childhood and adult blood samples can be identified in several ways. Some laboratories participate in proficiency testing programs operated by the College of American Pathologists (about 300 laboratories), the CDC (about 200 laboratories), and the state health departments of New York and Pennsylvania (about 70 laboratories each). About 250 of these laboratories ask the OSHA Salt Lake Technical Center to determine whether their performance meets OSHA requirements for biological monitoring.

The principal laboratories that test for lead and other toxins in biological samples can also be identified from a national survey by the Health Care Financing Administration. The survey is being conducted to support the Clinical Laboratory Improvement Amendments (CLIA) of 1988, which require certification for laboratories that test human materials for diagnosis or treatment.[21] It is expected that 200 000 to 250 000 facilities will be identified from the 640 000 potential facilities to which questionnaires have been sent. The questionnaire includes the number, type, and methodology of laboratory procedures and the qualifications of the supervisors and technicians. The CLIA also require certification for noncommercial laboratories (e.g., those operated within manufacturing industries), but these laboratories are not included in the survey.

Conclusion

New methods and additional resources are needed to achieve the national health objective of eliminating occupational lead exposures resulting in blood lead levels above 1.21 μmol/L.[1,22] Adult lead registries complement other forms of surveillance in identifying companies with excessive lead exposures. Reporting requirements and procedures should be standardized. Elevated blood lead levels should be a reportable condition nationwide, and a comprehensive national surveillance system should be established: clinical laboratories, regardless of location, should report cases to states with adult lead registries or, with those cases from states without registries, directly to the CDC-NIOSH registry. □

Acknowledgments

The author gratefully acknowledges the generous help of David Marion and Randy Young for computer programming; Dr Paul Seligman for consultation during manuscript preparation; Joseph Seta for information on the NOES; Dr W. Karl Sieber, Jr, for statistical consulting; staff members of the state registries; Dr Wayne Smith for information on the CLIA; and Dr James Melius.

References

1. Landrigan PJ, Silbergeld EK, Froines JR, Pfeffer RM. Lead in the modern workplace. *Am J Public Health.* 1990;80:907–908 (erratum 1991;81:162).
2. Rutstein DD, Mullan RJ, Frazier TM, Halperin WE, Melius JM, Sestito JP. Sentinel health events (occupational): a basis for physician recognition and public health surveillance. *Am J Public Health.* 1983;73:1054–1062.
3. Froines JR, Dellenbaugh CA, Wegman DH. Occupational health surveillance: a means to identify work-related risks. *Am J Public Health.* 1986;76:1089–1096.
4. Seligman PJ, Halperin WE, Mullan RJ, Frazier TM. Occupational lead poisoning in Ohio: surveillance using workers' compensation data. *Am J Public Health.* 1986;76:1299–1302.
5. Baker EL. Sentinel Event Notification System for Occupational Risks (SENSOR): the concept. *Am J Public Health.* 1989;79(suppl):18–20.
6. Froines J, Wegman D, Eisen E. Hazard surveillance in occupational disease. *Am J Public Health.* 1989;79(suppl):26–31.
7. Baser ME, Marion D. A statewide case registry for surveillance of occupational heavy metals absorption. *Am J Public Health.* 1990;80:162–164.
8. National Institute for Occupational Safety and Health. *National Occupational Exposure Survey.* Vols. 1–3. Washington, DC: US Dept of Health and Human Services; 1988, 1990. US Dept of Health and Human Services publications NIOSH 88–106, NIOSH 89–102, NIOSH 89–103.
9. Occupational Safety and Health Standards. Subpart Z: Toxic and hazardous substances—lead. Washington, DC: Office of the Federal Register, National Archives and Records Administration; 1985. 29 CFR § 1910.1025.
10. Froines JR, Baron S, Wegman D, O'Rourke S. Characterization of the airborne concentrations of lead in U.S. industry. *Am J Ind Med.* 1990;18:1–17.
11. Maizlish NA, Rudolph LA, Royce S, et al. *Elevated Blood Lead in California Adults, 1987–1990.* Berkeley, Calif: California Department of Health Services; 1991. Publication CDHS (COHP) SR90-001.

12. Office of Management and Budget. *Standard Industrial Classification Manual.* Washington, DC: US Government Printing Office; 1987.
13. Rudolph L, Sharp DS, Samuels S, Perkins C, Rosenberg J. Environmental and biological monitoring for lead exposure in California workplaces. *Am J Public Health.* 1990;80:921–925.
14. Goldman RH, Baker EL, Hannen M, Kamerow DB. Lead poisoning in automobile radiator mechanics. *N Engl J Med.* 1987; 317:214–218.
15. Safety and Health Regulations for Construction. Subpart J: Welding and cutting—welding, cutting, and heating in way of preservative coatings. Washington, DC: Office of the Federal Register, National Archives and Records Administration; 1988. 29 CFR § 1926.354.
16. Seligman PJ, Halperin WE. Targeting of workplace inspections for lead. *Am J Ind Med.* 1991;20:381–390.
17. Rappaport S. The rules of the game: an analysis of OSHA's enforcement strategy. *Am J Ind Med.* 1984;6:291–303.
18. Baker EL, Melius JM, Millar JD. Surveillance of occupational illness and injury in the United States: current perspectives and future directions. *J Public Health Policy.* 1988;9:198–221.
19. Maizlish N, Rudolph L, Sutton P, Jones JR, Kizer KW. Elevated blood lead in California adults, 1987: results of a statewide surveillance program based on laboratory reports. *Am J Public Health.* 1990;80:931–934.
20. Freund E, Seligman PJ, Chorba TL, Safford SK, Drachman JG, Hull HF. Mandatory reporting of occupational diseases by clinicians. *JAMA.* 1989;262:3041–3044.
21. Clinical Laboratory Improvement Amendments of 1988: Final rule (with comment). Washington, DC: Office of the Federal Register, National Archives and Records Administration; 1992. 42 CFR § 405 and others.
22. US Public Health Service. *Healthy people 2000: National Health Promotion and Disease Prevention Objectives.* Full report, with commentary. Washington, DC: US Dept of Health and Human Services; 1991. Publication PHS 91–50212.

Recent references related to this paper:

Tepper A. Surveillance of occupational lead exposure in New Jersey: 1986 to 1989. *Am J Public Health.* 1992;82:275–277.

Papanek PJ, Ward CE, Gilbert KM, Frangos SA. Occupational lead exposure in Los Angeles County: an occupational risk surveillance strategy. *Am J Ind Med.* 1992;21:199–208.

Call for Abstracts for Injury Control/Emergency Services Late-Breaker Session

The Injury Control and Emergency Health Services Section will again sponsor a "late-breaker" session during the APHA 1992 annual meeting in Washington, DC. The session will be held on Tuesday, November 10, and will feature work completed within the last few months (after the deadline for consideration in the regular symposia of the APHA annual meeting).

Submit abstracts of 250 words or fewer (any format) and a return envelope to Richard Waxweiler, Division of Injury Control, Centers for Disease Control, Mail Stop F-36, Atlanta, GA 30333; tel. (404) 488-4690.

Abstracts must be received by September 1, 1992. Decisions will be made by September 15, 1992.

Lower Respiratory Illness in Infants and Low Socioeconomic Status

Peter A. Margolis, MD, PhD, Robert A. Greenberg, MD, MSPH, Lynette L. Keyes, MS, MPH, Lisa M. LaVange, PhD, Robert S. Chapman, MD, MPH, Floyd W. Denny, MD, Karl E. Bauman, PhD, and Barbara W. Boat, PhD

ABSTRACT

Objectives. Infants from families of low socioeconomic status are said to suffer higher rates of lower respiratory illness, but this assertion has not been carefully examined.

Methods. We studied the frequency and determinants of lower respiratory illness in infants of different socioeconomic status (n = 393) by analyzing data from a community-based cohort study of respiratory illness during the first year of life in central North Carolina.

Results. The incidence of lower respiratory illness was 1.41 in the low socioeconomic group, 1.26 in the middle group, and 0.67 in the high group. The prevalence of persistent respiratory symptoms was 39% in infants in the low socioeconomic group, 24% in infants in the middle group, and 14% in infants in the high group. The odds of persistent respiratory symptoms in infants of low and middle socioeconomic status were reduced after controlling for environmental risk factors for lower respiratory illness. Enrollment in day care was associated with an increased risk of persistent symptoms among infants of high but not low socioeconomic status.

Conclusions. Infants of low socioeconomic status are at increased risk of persistent respiratory symptoms. This risk can be partly attributed to environmental exposures, most of which could be changed. (*Am J Public Health.* 1992;82:1119–1126)

Introduction

Lower respiratory illnesses are major causes of childhood morbidity in the United States. Approximately 30% of infants less than 1 year of age visit a doctor for a lower respiratory illness, and about 1% of infants who are healthy at birth are hospitalized for the illness.[1–6] Although the vast majority of acute lower respiratory illnesses are without complication, recent studies indicate that the occurrence of these illnesses in infancy may predispose to the development of obstructive airways disease in later life.[7–9]

Clinical experience suggests that infants of low socioeconomic status have higher rates of lower respiratory illness than more socially advantaged infants, but few studies have investigated this relationship. Previous studies of socioeconomic status and lower respiratory illness have provided limited information because they have examined only differences in mortality rates or because they have been hospital rather than community based.[5,6,10] The major studies of respiratory illness in childhood have focused on lower respiratory illness in middle-class families or on upper respiratory illness.[11,12] Thus, it is unclear whether the clinical impression that infants of low socioeconomic status have higher rates of lower respiratory illness is correct, and if so, whether it reflects an increased incidence of this illness, more severe illness, or more chronic illness.

The objectives of this study were to determine whether rates of acute lower respiratory illness and chronic respiratory symptoms differ among socioeconomic status groups and to identify environmental risk factors that might account for whatever differences in rates of disease were observed and that might be amenable to clinical or public health intervention.

Methods

Study Sample

We analyzed data from the control group of a community-based randomized trial designed to study the effects of an intervention to reduce infants' exposure to tobacco smoke and, thereby, their rate of lower respiratory illness in the first year of life. In the trial, an attempt was made to enroll a large proportion of the infants born in Alamance and Chatham counties in central North Carolina between 1986 and 1988. Recruitment of families took place at three hospitals at which 80% of the births in these counties occur. All mothers who gave birth to infants who weighed more than 2000 g and who did not require neonatal care outside the normal newborn nursery were asked to participate. At the time of recruitment, families

Peter A. Margolis, Robert A. Greenberg, and Floyd W. Denny are with the Department of Pediatrics; Lynette L. Keyes and Lisa M. LaVange are with the Frank Porter Graham Child Development Center; Karl E. Bauman is with the Department of Health Behavior, School of Public Health; and Barbara W. Boat is with the Department of Psychiatry, all at the University of North Carolina at Chapel Hill. Robert S. Chapman is with the US Environmental Protection Agency, Washington, DC.

Requests for reprints should be sent to Peter A. Margolis, MD, PhD, Cecil G. Sheps Center for Health Services Research, CB #7490, Chase Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7490.

This paper was submitted to the Journal July 16, 1991, and accepted with revisions December 10, 1991.

were told only that the study was intended to investigate infant health during the first year of life.

Data about potential risk factors for lower respiratory illness were collected during home interviews when the infants were 3 weeks and 12 months of age. Information was elicited by questionnaires administered to the infants' mothers and included demographic, environmental, and psychosocial characteristics. Data collectors were not told the objectives of the study.

Measurement of Respiratory Illness

We used two measures of respiratory illness: acute lower respiratory illness and persistent respiratory symptoms.

Symptoms of acute lower respiratory illness were measured by telephone calls to the families every 2 weeks. Telephone interviewers used standardized data collection forms and inquired about respiratory as well as nonrespiratory symptoms. Only the families with telephones, who constituted 85% of the study sample, were questioned about the presence of acute symptoms.

Acute lower respiratory illness was defined as the parents' report of the presence of cough, wheezing, and rattling in the chest. The incidence of acute lower respiratory illness was expressed as the number of episodes per child year at risk. The choice of this definition was based on a separate study of patients from the University of North Carolina Hospitals outpatient department and a private pediatric practice.[13] In this study, the combination of symptoms of cough plus wheezing and rattling in the chest had a sensitivity of 88% (95% CI = 0.80, 0.96) and a specificity of 68% (95% CI = 0.60, 0.76) in identifying lower respiratory illness (likelihood ratio = 2.75; 95% CI = 2.07, 3.65).

The presence of persistent respiratory symptoms was measured at the 12-month home interview using the American Thoracic Society children's questionnaire for studies of respiratory illness.[14] The questionnaire was modified slightly to make it more appropriate for use with infants. Infants who were reported to "usually cough" or "occasionally wheeze" were classified as having persistent respiratory symptoms. All families in the study were asked about the presence of persistent symptoms, regardless of whether they had telephones.

Measurement of Risk Factors for Respiratory Illness

The infants' socioeconomic status was classified into three levels according to the highest level of education achieved by the head of the household. Families in which the head of the household did not graduate from high school were classified as low socioeconomic status. If the head of household had completed only high school, the family was placed in the middle socioeconomic group; if more than a high school education, the high group. We used educational achievement because it appears to be a relatively stable measure of socioeconomic status in young families, whose incomes or occupational status may be increasing. Education is also the socioeconomic status measure that is most closely associated with health outcomes in children.[15,16]

Each infant's exposure to tobacco smoke was measured as the number of cigarettes smoked in the infant's presence during the week prior to the 12-month home visit. We chose to measure exposure at the 12-month home visit because other evidence from this study indicated that the amount of exposure remained stable throughout the entire first year of life.[17] In these analyses, infants were considered exposed to tobacco smoke if any smoking occurred in their presence during this time period. The questionnaire that was used for this measurement is described in detail elsewhere.[18]

Infants were considered exposed to day care if they were with at least one other child who was not their sibling for more than 4 hours per week and if they were cared for by someone other than their parents.[19,20] Day-care exposure was measured at the 12-month home interview. Household crowding was defined as the number of persons per room in the house and was classified into three levels: less than 0.5 persons per room, 0.5 through 1.0 persons per room, or more than 1.0 persons per room. Infants were considered exposed to smoke from a wood stove if the wood stove was the primary source of home heating.

Feeding and stress were measured at the 3-week home interview. Infants were classified as bottle fed if they received only formula and as breast fed if they received human milk with or without formula. Mothers' psychological stress was measured at the first home interview by seven questions from the Rand Mental Health Inventory.[21] The scores on the individual questions were averaged to provide a summary measure of stress. The summary measure was grouped into tertiles to define low, medium, and high levels of stress. We used this grouping to increase the clinical interpretability of our results.

A family history of allergy was defined as the presence of asthma or hay fever in either parent. The season of an infant's birth was considered to be respiratory if the infant was born between October and March and nonrespiratory if the infant was born between April and September.[22]

Statistical Analyses

The variance of the incidence density of lower respiratory illness was computed using the Taylor series method.[23] The significance of the differences in incidence of lower respiratory illness across socioeconomic status groups was assessed by a test for linear trend using a Wald χ^2 statistic.[24]

Relationships between the respiratory risk factors, socioeconomic status, and persistent respiratory symptoms were summarized by prevalence relative risks, and statistical significance was assessed by χ^2 tests. After calculating the unadjusted relative risk of persistent respiratory symptoms for the low and middle socioeconomic groups compared with the high socioeconomic group, Cochran-Mantel-Haenzel methods were used to adjust the relative risk estimates and confidence intervals for the effects of individual risk factors.[25]

We used logistic regression to analyze to what extent the relationship between socioeconomic status and persistent respiratory symptoms could be accounted for by simultaneously considering interactions between socioeconomic status and other risk factors for lower respiratory illness and confounding by other risk factors.[26,27] In the models, only the risk factors associated with both the presence of persistent symptoms and socioeconomic status were considered.[27] The predictor variables in the full model were examined for evidence of collinearity using standard analysis techniques.[28] Crowding, stress, and socioeconomic status were included in the logistic models as interval variables after plots confirmed that each had a linear relationship with the outcome variable on a logit scale.[29]

To determine whether effect modification was present, the full model, which included all two-way interactions between the risk factors and socioeconomic status, was simplified by deleting statistically

nonsignificant interactions in a backward selection procedure (maximum likelihood χ^2 statistic; P value = .1 to remain in the model). Only two-way interaction terms were considered for inclusion in the models, and the sole significant interaction was between socioeconomic status and day care. After assessing for effect modification, we determined whether any of the risk factors could be removed from the model including the socioeconomic status and day-care interaction without an important change in the estimated effect of socioeconomic status.[27] The final model was used to estimate the prevalence of persistent symptoms for groups of infants given various risk factors.[25]

Results

Characteristics of Study Participants

There were 2332 births in which the infant was eligible to participate in the intervention trial. Of the eligible families, 1091 (47%) agreed to enroll. Of the 516 infants who were randomized to the control group, 396 (77%) completed the study and were included in the analysis for this paper. Three infants were excluded because their race was neither White nor Black. The socioeconomic status of the 393 infants studied was as follows: 20% low, 31% middle, and 49% high.

Compared with the 1241 infants from families refusing enrollment, study infants were more likely to be of high socioeconomic status and were more often Black (Table 1). Study infants were less likely to have mothers who smoked. In families in which mothers did smoke, however, study mothers and mothers from families refusing enrollment smoked the same average number of cigarettes per day. The 33 families without telephones, for whom we do not have a measurement of the incidence of acute lower respiratory illness, were more often of low or middle socioeconomic status than those in the study group as a whole. The distribution of socioeconomic status among those without telephones was as follows: 49% low, 42% middle, and 9% high. However, the distribution of other risk factors in these groups appeared similar to the distribution of risk factors in the entire sample.

When compared with the infants of the 120 families who enrolled but did not complete the study, infants of families who completed the study were more often of high socioeconomic status (49% vs 20%) but less often Black (30% vs 47%) and less likely to have mothers who smoked (21% vs 37%). The distribution among socioeconomic groups of potential risk factors for lower respiratory illness (household crowding, exposure to tobacco smoke) was the same among study infants and dropouts.

TABLE 1—Characteristics of Study Families and Families Refusing Enrollment

	Study Families, (n = 393)	Families Refusing Enrollment, (n = 1241)
Socioeconomic status		
Low	20	26
Middle	31	43
High	49	31
Black	30	23
Mother smoked	21	27
Mean no. of cigarettes mother smoked per day[a]	15	15

[a]Among mothers who smoked.

TABLE 2—Prevalence of Persistent Respiratory Symptoms in Socioeconomic Groups

Socioeconomic Status	n	Persistent Symptoms, %	Relative Risk (95% CI)
Low	79	39	2.9 (1.9, 4.5)
Middle	122	24	1.8 (1.1, 2.8)
High	192	14	1.0

TABLE 3—Relationship of Socioeconomic Status to Prevalence of Risk Factors for Respiratory Illness

	Socioeconomic Status, %		
Risk Factor	Low (n = 79)	Middle (n = 122)	High (n = 192)
Demographic			
Male	48	48	49
Black	49	40	16
Birthweight 2000–2499 g	10	5	2
Born in nonrespiratory season	43	57	52
Familial			
Family history of allergy	20	24	34
Family history of respiratory disease	15	16	18
Maternal stress (high)	51	37	16
Environmental			
Exposed to tobacco smoke	79	78	35
Household crowding (> 0.5 persons/room)	70	61	29
Wood stove	14	7	6
Gas cooking	17	13	6
Day care	28	44	52
Bottle fed	82	73	33

Frequency of Respiratory Illness

Both the incidence of acute lower respiratory illness and the prevalence of persistent symptoms had dose-response relationships with socioeconomic status. The incidence of acute lower respiratory illness (episodes per child year) was 1.41 for low socioeconomic infants, 1.26 for middle socioeconomic infants, and 0.67 for high socioeconomic infants (P = .009 for linear trend). The relationship between the prevalence of persistent respiratory symptoms and socioeconomic status was even more marked (Table 2). Among in-

TABLE 4—Relative Risk of Persistent Respiratory Symptoms among Infants Exposed to Risk Factors

Risk Factor	Relative Risk	95% CI	P
Demographic			
Male	1.06	0.73, 1.55	NS
Black	1.75	1.21, 2.56	.003
Birthweight 2000–2499 g	1.85	0.94, 3.62	NS
Born in nonrespiratory season	0.81	0.56, 1.18	NS
Familial			
Family history of allergy	1.26	0.81, 1.96	NS
Family history of respiratory disease	1.35	0.82, 2.24	NS
Maternal stress (high)	1.59	1.01, 2.49	0.45
Environmental			
Exposed to tobacco smoke	1.85	1.23, 2.77	.003
Household crowding (> 0.5 persons/room)	2.53	1.39, 3.21	<.001
Wood stove	1.25	0.65, 2.35	NS
Gas cooking	1.12	0.63, 2.04	NS
Bottle fed	2.07	1.32, 3.24	.001

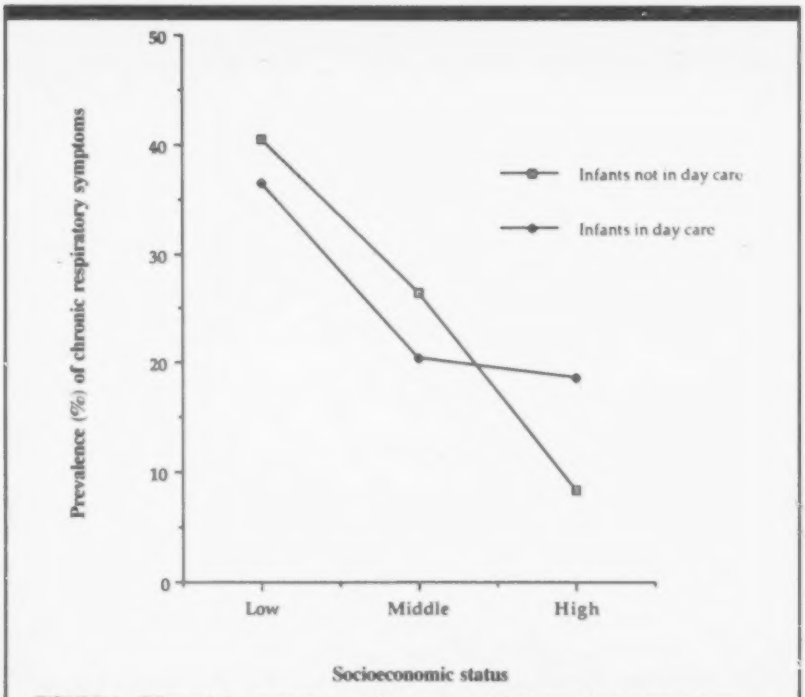


FIGURE 1—Effect of day care on prevalence of persistent respiratory symptoms.

fants of low socioeconomic status, 39% had persistent symptoms; only 14% had persistent symptoms in the high socioeconomic group. Infants in the low socioeconomic group were 2.9 (95% CI = 1.9, 4.5) times more likely than infants of high socioeconomic status to have persistent respiratory symptoms.

Infants who had persistent respiratory symptoms were more likely to have had an acute lower respiratory illness than those who did not have persistent symptoms. Of infants with persistent symptoms, 68% had at least one episode of acute lower respiratory illness, whereas only 32% of infants without persistent symptoms had any episodes of the illness ($P < .0001$, χ^2). However, the presence of persistent symptoms did not represent repeated episodes of acute lower respiratory illness. Among infants with persistent symptoms who had experienced an acute lower respiratory illness, most (63%) had only one or two episodes, and only 15% had more than three episodes.

Relationship of Risk Factors to Socioeconomic Status and Persistent Respiratory Symptoms

To investigate how much of the effect of socioeconomic status could be related to demographic, environmental, or psychosocial factors, we evaluated its association with persistent respiratory symptoms after controlling for risk factors for lower respiratory illness[30–37] that were also associated with both socioeconomic status and persistent respiratory symptoms in our data. We focused on persistent or continuing respiratory symptoms because they may be indicative of the long-term impact of acute respiratory illness. In these analyses, factors that contributed to the effect of socioeconomic status would lead to a decrease in the association between socioeconomic status and persistent respiratory symptoms.

Table 3 shows the relationship of the risk factors we studied to socioeconomic status, and Table 4 indicates their relationship to persistent respiratory symptoms. Of the demographic, familial, and environmental factors, only the infants' sex and season of birth were not associated with socioeconomic status. Although the association of low birthweight with persistent symptoms was not statistically significant ($P = .074$), the prevalence of this risk factor in the study sample was low (4.6%) because infants weighing less than 2000 g were excluded from the study. We decided to adjust for low birthweight in studying the relationship between socioeconomic status and persistent symptoms, however, because the magnitude of its relationship was clinically relevant (relative risk = 1.85) and because there was previous evidence of its association with persistent respiratory illness.[38] Thus, household crowding, exposure to tobacco smoke, style of feeding, birthweight, stress, and race were related to both socioeconomic status and persistent respiratory symptoms.

The effect of day care on persistent symptoms differed significantly according to the socioeconomic status of the infant (Figure 1). Among infants of high socioeconomic status, enrollment in day care was associated with an increase in the prevalence of persistent symptoms. In infants of low and middle socioeconomic status, however, enrollment in day care was associated with a reduction in the prevalence of persistent symptoms.

Effect of Risk Factors on the Relationship of Socioeconomic Status to Persistent Respiratory Symptoms

To evaluate the effect of the risk factors on the relationship between socioeconomic status and persistent respiratory symptoms, we compared the unadjusted relative risk of persistent symptoms in the low and middle socioeconomic groups with the relative risk of persistent symptoms after adjusting for each risk factor separately by stratified analysis (Table 5). The greatest reduction in the relative risk occurred after control for exposure to household crowding and tobacco smoke among infants of low socioeconomic status. Control for crowding decreased the relative risk among infants of low compared with high socioeconomic status from 2.9 to 2.2, and control for tobacco smoke exposure reduced the relative risk of persistent symptoms to 2.3.

We used logistic regression to evaluate how much of the effect of socioeconomic status on persistent symptoms might be produced by all the risk factors together. We compared the odds of persistent symptoms in a model that did not adjust for the effect of the risk factors with the odds of persistent symptoms in a model that adjusted for all the risk factors together (see Appendix). The odds ratios estimated from the models do not approximate the relative risks calculated by stratified analysis because of the high prevalence of persistent symptoms.

In the model that did not adjust for the effect of the risk factors, the odds ratio of persistent symptoms among infants of low compared with high socioeconomic status was 12.4 (95% CI = 7.5, 20.3) for those not in day care (Table 6). Among infants in day care, the odds ratio of persistent symptoms was 2.8 (95% CI = 1.6, 4.8). Control for all the risk factors reduced the odds of persistent symptoms among infants of low compared with high socioeconomic status from 12.4 to 6.9 (95% CI = 3.8, 12.2) for infants not in day care and from 2.8 to 1.3 (95% CI = 0.7, 2.4) for infants in day care. After accounting for all the risk factors, the effect of socioeconomic status remained significant only for infants not in day care.

To illustrate how much these risk factors contributed to the effect of socioeconomic status on persistent symptoms, we estimated the prevalences of persistent respiratory symptoms from the logistic model for infants of different socioeconomic status groups who were exposed to

TABLE 5—Effect of Socioeconomic Status on the Risk of Persistent Respiratory Symptoms

	Relative Risk (95% CI)		
	Low Socioeconomic Status	Middle Socioeconomic Status	High Socioeconomic Status[a]
Unadjusted	2.9 (1.9, 4.5)	1.8 (1.1, 2.8)	1.0
Adjusted			
Stress	3.5 (2.1, 5.7)	2.3 (1.3, 3.8)	1.0
Feeding	3.1 (1.8, 5.3)	2.1 (1.2, 3.7)	1.0
Birthweight	2.8 (1.8, 4.4)	1.7 (1.1, 2.8)	1.0
Race	2.6 (1.6, 4.2)	1.6 (0.99, 2.7)	1.0
Exposure to tobacco smoke	2.3 (1.5, 3.7)	1.5 (0.89, 2.4)	1.0
Crowding	2.2 (1.4, 3.6)	1.6 (0.99, 2.5)	1.0

[a]Reference group

TABLE 6—Unadjusted and Adjusted Odds of Persistent Respiratory Symptoms in Socioeconomic Groups

	Odds Ratio (95% CI)		
	Low Socioeconomic Status	Middle Socioeconomic Status	High Socioeconomic Status[a]
Infants not in day care			
Unadjusted	12.4 (7.5, 20.3)	3.5 (2.1, 5.8)	1.0
Adjusted[b]	6.9 (3.9, 12.2)	2.6 (1.5, 4.6)	1.0
Infants in day care			
Unadjusted	2.8 (1.6, 4.8)	1.7 (1.0, 2.9)	1.0
Adjusted[b]	1.3 (0.7, 2.4)	1.2 (0.6, 2.1)	1.0

[a]Reference group.
[b]Adjusted for crowding, tobacco smoke, bottle feeding, low birthweight, stress, Black race.

TABLE 7—Estimated Prevalence of Persistent Symptoms Due to Exposure to Risk Factors among Infants Not in Day Care

	Socioeconomic Status		
Risk Factor	Low	Middle	High
None	.22	.09	.04
Tobacco smoke alone	.28	.13	.05
Tobacco smoke, crowding	.59	.28	.15
All risk factors[a]	.80	.61	.37

[a]Risk factors: crowding, tobacco smoke, bottle feeding, low birthweight, stress, Black race.

TABLE 8—Estimated Prevalence of Persistent Symptoms Due to Exposure to Risk Factors among Infants in Day Care

	Socioeconomic Status		
Risk Factor	Low	Middle	High
None	.12	.11	.10
Tobacco smoke alone	.17	.15	.13
Tobacco smoke, crowding	.42	.39	.35
All risk factors[a]	.67	.64	.61

[a]Risk factors: crowding, tobacco smoke, bottle feeding, low birthweight, stress, Black race.

various risk factors. Tables 7 and 8 show the estimated effect of the risk factors among infants in day care (44% of the sample) and not in day care. For example, for infants of low socioeconomic status not in day care (Table 7) and not exposed to any risk factors, the prevalence of persistent symptoms would be about 22%. Exposure to tobacco smoke would increase the prevalence of persistent symptoms to 28%; crowding and smoking to 59%. Exposure to all the risk factors would increase the prevalence of persistent symptoms to 80%. Although these risk factors increased the probability of persistent respiratory symptoms in all socioeconomic groups, the effect of the risk factors on the prevalence of symptoms would be greatest in the low socioeconomic group. Among infants of low socioeconomic status, the difference in prevalence between those exposed to all the risk factors and those exposed to none would be 58%. Among infants of high socioeconomic status, however, the difference would be only 33%. For infants who were enrolled in day care, the effects of exposure to risk factors were similar in all three socioeconomic groups (Table 8).

Discussion

Our results indicate that infants of low socioeconomic status experience a higher incidence of acute lower respiratory illness and a higher prevalence of persistent respiratory symptoms than infants of high socioeconomic status. These results could have occurred if families of low socioeconomic background were more likely to report symptoms in their children. However, there is evidence that highly educated families are more likely than less-educated families to report respiratory symptoms.[39] In our assessment of the validity of our acute respiratory symptoms questionnaire, we did not detect clinically or statistically significant differences by socioeconomic status in responses to the questions, but our sample was small.

The differences we observed in the frequencies of respiratory symptoms across socioeconomic strata are also not explained by how we measured respiratory illness. Inaccuracy in the measurement of acute lower respiratory illness and persistent respiratory symptoms would have led to random misclassification of respiratory illness and diminished the observed differences between socioeconomic groups. Therefore, the effects of socioeconomic status we found are likely to be conservative estimates.

The large number of infants who were not in the study sample may limit the generalizability of our results. Only about half of the eligible families were enrolled in our study, and there were some differences in the distribution of risk factors and socioeconomic status between the study population and the target population. For example, infants in the study sample were of higher socioeconomic status, and fewer had mothers who smoked. This exposure to fewer risk factors among study infants compared with the population as a whole also indicates that our results may represent conservative estimates of the differences between socioeconomic groups.

Although socioeconomic status affects both the incidence of acute lower respiratory illness and the prevalence of persistent respiratory symptoms, its impact appears to be greater on the prevalence of persistent symptoms. One possible explanation for the differing effects of socioeconomic status on acute lower respiratory illness and persistent symptoms may be that factors related to socioeconomic status have a more pronounced effect on infants' risk of developing serious or chronic lower respiratory illness than on their risk of becoming infected. This explanation is consistent with the results of other studies that indicate that socioeconomic status appears to have little effect on rates of upper respiratory illness, but more of an effect on acute lower respiratory illness and a still greater influence on serious and chronic forms of this illness.[12,40]

There are insufficient data to assess the long-term clinical significance of persistent respiratory symptoms in 1-year-old infants. In studies of older children, however, a history of chronic cough and occasional wheezing has been associated with decrements in measurements of pulmonary function,[41] and such reduced ventilatory function may be associated with chronic obstructive airways disease.[7]

There is little direct information about differences in incidence of acute lower respiratory illness across socioeconomic groups. However, the one previous study of this question found differences in incidence rates that were similar to ours. Gardner et al.[42] described the incidence of lower respiratory illness in 23 low socioeconomic infants and 48 high socioeconomic infants who were participants in a community viral surveillance project in Houston. Lower respiratory illness was diagnosed by physical examination at home visits every 1 to 2 weeks. Although we used a telephone survey to identify illnesses in a larger and community-based cohort identified at birth, the incidence rates of lower respiratory illness they observed (1.43 episodes per child year in the low socioeconomic group, 1.0 episodes per child year in the middle socioeconomic group, and 0.72 episodes per child year in the high socioeconomic group) were similar to ours.

Previous evidence suggests that there are large differences in the severity of lower respiratory illness in infants of low compared with high socioeconomic status. For example, infants of low socioeconomic status who are healthy at birth have mortality rates from lower respiratory illness in the first year of life that are 2 to 4 times higher than those found among more advantaged children,[10,43] and other studies of infants suggest that hospitalization rates for lower respiratory illness are 5 to 10 times higher among poor as compared with middle-class infants.[5,44]

Chronic respiratory symptoms, too, are more prevalent in school-aged children of low socioeconomic status, but the effects of socioeconomic status on chronic respiratory symptoms in infancy have not been studied longitudinally. For example, Colley and Reid found the prevalence of chronic cough to be about three times higher in school children of low compared with high socioeconomic status.[45] Studies by Ware et al. and Schenker et al. also found that chronic cough and a history of serious "chest illness" in infancy were more common in children of low socioeconomic status.[46,47]

Of the risk factors for respiratory illness we found to be important in our study, most are supported by previous work.[30-37] The finding that day care was associated with a somewhat reduced prevalence of persistent symptoms in low socioeconomic families and an increased prevalence in high socioeconomic families was unexpected and has not been previously reported. It deserves exploration in another study. It is possible that the protective effect of day care among the infants of low and middle socioeconomic status could be related to the smaller size of their day-care groups or to a reduction in these infants' exposure to environmental factors such as tobacco smoke.[48] We found that families of low socioeconomic status tended to use day care less often and that infants in these families were more often cared for alone or in small groups (< 5 children), whereas infants of high socioeconomic status were more

commonly in large-group day-care settings in which the likelihood of exposure to infection is greater.

Our results indicate that much of the effect of socioeconomic status can be explained by factors that are related to the infant's home environment, most of which are mutable (exposure to tobacco smoke, crowding, stress, low birthweight, and bottle feeding). In addition, exposure to risk factors for respiratory illness is not only more common among families of low socioeconomic status, but the effect of the exposure also seems to be the greatest in this group, at least among infants not in day care. From a public health perspective, these findings indicate that infants of low socioeconomic status are especially likely to benefit from interventions that reduce their exposure to risk factors for respiratory illness. Given the high prevalence of persistent symptoms in infants of low socioeconomic status, even small changes in what are potentially mutable environmental risk factors could have a substantial impact on the health of poor infants.

Finally, our results indicate that a substantial portion of the effect of socioeconomic status cannot be explained by known environmental exposures. Thus, we should continue to study those at highest risk in order to understand better the factors that subject the socially disadvantaged to excessive burdens of illness. It may be necessary to develop interventions aimed not only at changing exposure to individual risk factors, but also at changing the social, political, and economic environment that promotes the effects of low socioeconomic status. □

Acknowledgments

This research was supported by Grant 28895 from the National Heart, Lung, and Blood Institute, National Institutes of Health; the Robert Wood Johnson Clinical Scholars Program; and the Primary Care Research Training Program. The study was approved by the Committee on Research Involving Human Subjects of the University of North Carolina School of Medicine.

We are grateful to the pediatricians of Alamance and Chatham counties, North Carolina, for supporting our study of their patients. The cooperation of the staffs of Alamance County Hospital, Alamance Memorial Hospital, and North Carolina Memorial Hospital is also greatly appreciated. We thank the staff of the Infant Health Study, whose work made this study possible. We also wish to thank Dr. Robert Fletcher, Codirector of the Robert Wood Johnson Clinical Scholars Program, for encouraging us to pursue this topic and for his guidance while Dr. Margolis conducted this research project as a Robert Wood Johnson Clinical Scholar.

The views expressed are those of the authors and do not necessarily represent those of the Robert Wood Johnson Foundation.

References

1. Glezen WP, Loda FA, Clyde WA Jr, et al. Epidemiologic patterns of acute lower respiratory disease of children in a pediatric group practice. *J Pediatr.* 1971;78:397–406.
2. Henderson FW, Clyde WA Jr, Collier A, et al. The etiologic and epidemiologic spectrum of bronchiolitis in pediatric practice. *J Pediatr.* 1979;95:183–190.
3. Foy HM, Cooney MK, Maletzky AJ, Grayston JT. Incidence and etiology of pneumonia, croup and bronchiolitis in preschool children belonging to a prepaid medical care group over a four-year period. *Am J Epidemiol.* 1973;97:80–92.
4. Maletzky AJ, Cooney MK, Luce R, Kenny GE, Grayston JT. Epidemiology of viral and mycoplasmal agents associated with childhood lower respiratory illness in a civilian population. *J Pediatr.* 1971;78:407–414.
5. Kim HW, Arrobio JO, Brandt CD, et al. Epidemiology of respiratory syncytial virus infection in Washington, D.C. I. Importance of the virus in different respiratory tract disease syndromes and distribution of infection. *Am J Epidemiol.* 1973;98:216–225.
6. Glezen WP, Denny FW. Epidemiology of acute lower respiratory disease in children. *N Engl J Med.* 1973;288:498–505.
7. Samet JM, Tager IB, Speizer FE. The relationship between respiratory illness in childhood and chronic air-flow obstruction in adulthood. *Am Rev Respir Dis.* 1983; 127:508–523.
8. Woolcock AJ, Leeder SR, Peat JK, Blackburn CRB. The influence of lower respiratory illness in infancy and childhood and subsequent cigarette smoking on lung function in Sydney school children. *Am Rev Respir Dis.* 1979;120:5–14.
9. McConnochie KM, Roghmann KJ. Bronchiolitis as a possible cause of wheezing in childhood: new evidence. *Pediatrics.* 1984; 74:1–10.
10. Spurlock CW, Hines MW, Skagg JW, Hernandez CE. Infant death rates among the poor and non-poor in Kentucky, 1982 to 1983. *Pediatrics.* 1987;80:262–269.
11. Dingle JH, Badger GF, Jordan WS Jr. *Illness in the Home: A Study of 25 000 Illnesses in a Group of Cleveland Families.* Cleveland, Ohio: The Press of Western Reserve University, 1964.
12. Monto AS, Ullman BM. Acute respiratory illness in an American community: the Tecumseh study. *JAMA.* 1974;227:164–169.
13. Margolis P. [The value of clinical history in detecting lower respiratory illness in infants]. Unpublished raw data; 1989.
14. Ferris BG. Epidemiology standardization project. *J Am Thorac Soc.* 1978;118:1–119.
15. Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev.* 1988;10:87–121.
16. Cleland J, Van Ginneken J. Maternal schooling and childhood mortality. *J Biosoc Sci.* 1989;10(suppl):13–34.
17. Greenberg RA, Bauman KE, Strecher VJ, et al. Passive smoking in the first year of life. *Am J Public Health.* In press.
18. Bauman KE, Strecher VJ, Greenberg RA, Haley NJ. A comparison of biochemical and interview measures of the exposure of infants to environmental tobacco smoke. *Eval Health Professions.* 1989;12:179–191.
19. Anderson LJ, Parker RA, Strikas RA, et al. Day-care center attendance and hospitalization for lower respiratory tract illness. *Pediatrics.* 1988;82:300–308.
20. Fleming DW, Cochi SL, Hightower AW, Broome CV. Childhood upper respiratory tract infections: to what degree is incidence affected by day-care attendance? *Pediatrics.* 1987;79:55–60.
21. Ware JE Jr, Johnson SA, Davies-Avery A, Brook RH. *The Conceptualization and Measurement of Health for Adults in the Health Insurance Study: Volume III, Mental Health.* Santa Monica, Calif: Rand Corporation; 1978. R1987/3-HEW.
22. Denny FW, Clyde WA Jr. Acute lower respiratory tract infections in nonhospitalized children. *J Pediatr.* 1986;108:635–646.
23. Cochran WG. *Sampling Techniques.* New York, NY: Wiley; 1977:169–171.
24. Bishop YMM, Feinberg SE, Holland PW. *Discrete Multivariate Analysis: Theory and Practice.* Cambridge, Mass: The MIT Press; 1975.
25. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods.* Belmont, Calif: Lifetime Learning Publishers; 1982.
26. Koch G, Edwards S. Logistic regression. In: Johnson NL, Kotz SP, eds. *Encyclopedia of Statistics.* New York, NY: Wiley; 1983:128–133.
27. Greenland S. Modeling and variable selection in epidemiologic analysis. *Am J Public Health.* 1989;79:340–349.
28. Belsley DA, Kuh E, Welsch RE. Detecting and assessing collinearity. In: *Regression Diagnostics: Identifying Influential Data.* New York, NY: Wiley, 1980:85–160.
29. Harrell FE, Lee KL. The practical value of logistic regression. In: *Proceedings of the Tenth Annual SUGI Conference.* Cary, NC: SAS Users Group International; 1985: 1031–1036.
30. Samet JM, Marbury MC, Spengler JD. Health effects and sources of indoor air pollution. Part I. *Am Rev Respir Dis.* 1987; 136:1486–1508.
31. Samet JM, Marbury MC, Spengler JD. Health effects and sources of indoor air pollution. Part II. *Am Rev Respir Dis.* 1988; 137:221–242.
32. Wald ER, Dashefsky B, Byers C, Taylor F. Frequency and severity of infections in day care. *J Pediatr.* 1988;112:540–545.
33. Kovar MG, Serdula MK, Marks JS, Fraser DW. Review of the epidemiologic evidence for an association between infant feeding and infant health. *Pediatrics.* 1984; 74(suppl):615–638.
34. Fergusson DM, Horwood LJ, Shannon FT, Taylor B. Breast-feeding, gastrointestinal and lower respiratory illness in the first two years. *Aust Pediatr J.* 1981;17:191–195.
35. Watkins CJ, Leeder SR, Corkhill RT. The relationship between breast and bottle feeding and respiratory illness in the first

year of life. *J Epidemiol Community Health.* 1979;33:180–182.
36. Boyce WT, Jensen EW, Cassel JC, Collier AM, Smith AH, Ramey CT. Influence of life events and family routines on childhood respiratory tract illness. *Pediatrics.* 1977; 60:609–615.
37. Graham NMH, Douglas RM, Ryan P. Stress and acute respiratory infection. *Am J Epidemiol.* 1985;124:389–401.
38. Jason JM. Infectious disease-related deaths of low birth weight infants, United States, 1968 to 1982. *Pediatrics.* 1988;81: 58–62.
39. Tsui AO, DeClerque J, Mangani N. Maternal and sociodemographic correlates of child morbidity in Bas Zaire: the effects of maternal reporting. *Soc Sci Med.* 1988;26: 701–713.
40. Graham NMH. The epidemiology of acute respiratory infections in children and adults: a global perspective. *Epidemiol Rev.* 1990;12:149–178.
41. Mitchell C, Miles J. Lower respiratory tract symptoms in Queensland schoolchildren. The questionnaire: its reliability and validity. *Aust N Z J Med.* 1983;13:264–269.
42. Gardner G, Frank A, Taber LH. Effects of social and family factors on viral respiratory infection and illness in the first year of life. *J Epidemiol Community Health.* 1984; 38:42–48.
43. Druschel CM, Hale CB. Postneonatal mortality among normal birth weight infants in Alabama, 1908 to 1983. *Pediatrics.* 1987; 80:869–872.
44. Glezen WP, Paredes A, Allison JE, Taber LH, Frank AL. Risk of respiratory syncytial virus infection for infants from low-income families in relationship to age, sex, ethnic group, and maternal antibody level. *J Pediatr.* 1981;98:708–715.
45. Colley JRT, Reid DD. Urban and social origins of childhood bronchitis in England and Wales. *Br Med J.* 1970;2:213–217.
46. Ware JE Jr, Dockery DW, Spiro A, Speizer FE, Ferris BG. Passive smoking, gas cooking, and respiratory health of children living in six cities. *Am Rev Respir Dis.* 1984;129: 366–374.
47. Schenker MD, Samet JM, Speizer FE. Risk factors for childhood respiratory disease. The effect of host factors and home environmental exposures. *Am Rev Respir Dis.* 1983;128:1038–1043.
48. Wright AL, Holberg C, Martinez FD, Taussig LM, Group Health Medical Associates. Relationship of parental smoking to wheezing and nonwheezing lower respiratory tract illnesses in infancy. *J Pediatr.* 1991;118:207–214.

APPENDIX—Estimated Logistic Regression Model Parameters for Estimation of Prevalence of Persistent Respiratory Symptoms

Effect	β	SE	*P*
Intercept	−3.38	.548	.0001
Socioeconomic status	.967	.289	.0008
Crowding	.650	.265	.0141
Tobacco smoke	.347	.328	.291
Feeding	.099	.354	.779
Birthweight	.351	.636	.581
Race	.282	.329	.392
Stress	.155	.178	.383
Day care	.975	.494	.048
Day care × socioeconomic status	−.826	.398	.038

Note. The logistic model had the following form: logit P (persistent symptoms) = −3.38 + .967 (socioeconomic status) + .650 (crowding) + .347 (tobacco smoke) + .099 (feeding) + .351 (birthweight) + .155 (stress) + .282 (race) + .975 (day care) − .826 (day care × socioeconomic status). Exposure to tobacco smoke, bottle feeding and day care, and Black race and low birthweight were coded as dichotomous variables, with the presence of the variable coded as 1. Socioeconomic status (low, middle, high) and crowding (< .5, .5–1, > 1) were coded as three level ordinal variables, with low socioeconomic status and the most crowded living conditions coded as 2. Stress (low, medium, high) was coded into three levels, with the highest level coded as 3.

Call for Papers for Social Policy Journal

Research in Social Policy publishes theoretically informed interdisciplinary analyses of the values, attitudes, beliefs, politics, and organizational contexts of policy makers and their impact on target populations and problems. Critical theoretical and methodological perspectives are encouraged. December 1, 1992, is the deadline for volume 4, "Science and Technology Policy Studies," and June 15, 1993, is the deadline for volume 5, "Medical and Health Policies." Manuscripts should not exceed 40 pages, including notes. Send papers and inquiries to John H. Stanfield II, Department of Sociology, College of William and Mary, Williamsburg, VA 23187-8795.

Imported Leprosy in the United States, 1978 through 1988: An Epidemic without Secondary Transmission

Timothy D. Mastro, MD, Stephen C. Redd, MD, and Robert F. Breiman, MD

ABSTRACT

Objectives. Leprosy remains a major health problem in many regions of the world. In the United States, although leprosy continues to be reported, approximately 90% of cases are imported (i.e., occur among immigrants and refugees). An increase in imported cases began in 1978. This study was conducted to analyze this trend and to characterize the contributing cases.

Methods. Centers for Disease Control leprosy surveillance data from 1971 through 1988 were analyzed.

Results. The number of imported cases reported annually was relatively constant from 1971 through 1977 (mean = 119 per year), increased to 307 in 1985, and then decreased to 102 in 1988. Of the 957 excess cases reported from 1978 through 1988, 73.4% were among persons from Southeast Asia, including 51.3% from Vietnam, Cambodia, and Laos (Indochina). There was no coincident increase in indigenous cases of leprosy; the mean annual number of such cases was 17.7 (range = 10 to 29). Leprosy remains endemic in Texas, Hawaii, Louisiana, and possibly California.

Conclusions. An epidemic of imported leprosy began in the United States in 1978, peaked in 1985, and ended by 1988. This increase was primarily due to cases among refugees from Indochina and was limited by a decrease in the influx of Indochinese refugees in the mid-1980s. There is no evidence that these cases resulted in transmission in the United States. (*Am J Public Health.* 1992:82:1127–1130)

Introduction

Leprosy (Hansen disease) is a major health problem in many regions of the world, affecting more than 11 million persons in the 93 countries where it remains endemic.[1,2] The burden of leprosy is greatest in South and Southeast Asia, Africa, and Latin America.[3] This chronic, debilitating, infectious bacterial disease, caused by *Mycobacterium leprae,* has historically been associated with social stigmatization; before the advent of modern chemotherapy, leprosy patients were traditionally sequestered.[4] In northern Europe, Hawaii, Japan, and the continental United States, the incidence of leprosy had declined well before the introduction of effective chemotherapy in the mid-20th century.[1,5] In the United States, although cases of leprosy continue to be reported, approximately 90% of these cases involve persons entering the country from endemic areas and are defined as imported.[6,7] Leprosy has remained endemic in the United States in Texas, Hawaii, Louisiana, and California.[7] An increase in imported cases starting in 1978 was identified by review of Centers for Disease Control (CDC) surveillance data from 1971 through 1981.[7] We analyzed CDC surveillance data through 1988 to determine whether this trend continued and to characterize the contributing cases. Previously reported cases from 1971 through 1981 were included in the analysis to allow for the evaluation of trends over nearly two decades.[7,8]

Methods

Cases of leprosy are reported to CDC by state and territorial health departments using a standardized report form. Diagnosis is determined by clinical and laboratory (biopsy and/or skin smear) evaluation; type of leprosy is reported as lepromatous, dimorphous/borderline, tuberculoid, or indeterminate. The report includes patient's date of birth, date of entry into the United States, date of onset of symptoms of leprosy, date leprosy was first diagnosed by a physician, and date of report. Cases are categorized as indigenous or imported. Indigenous cases are defined as illness in persons born in the United States who have neither lived nor been stationed for military service outside the country. Imported cases are defined as illness in persons born outside the United States with onset either before or after entering the country, or illness in persons born in the United States who traveled outside the country and may have acquired infection in leprosy-endemic areas. Reports were excluded from the analysis if missing data prevented categorization as imported or indigenous.

Cases were recorded by the year of report, which often differed from the year of onset. Data on immigrant and refugee populations entering the United States were obtained from the US Immigration and Naturalization Service and the Department of State, Bureau for Refugee Programs, respectively. Asian countries were divided into three regions: (1) Southeast Asia—Cambodia, Indonesia, Laos, Malaysia, Myanmar (Burma), Philippines, Singapore, Thailand, and Vietnam; (2) East Asia—China, Hong Kong, Japan, Korea, and Taiwan; and (3) South and

The authors are with the Respiratory Diseases Branch, Division of Bacterial and Mycotic Diseases, National Center for Infectious Diseases, Centers for Disease Control, Atlanta, Ga.

Requests for reprints should be sent to Timothy D. Mastro, MD, Mailstop C-09, Centers for Disease Control, Atlanta, GA 30333.

This paper was submitted to the Journal September 30, 1991, and accepted with revisions December 31, 1991.

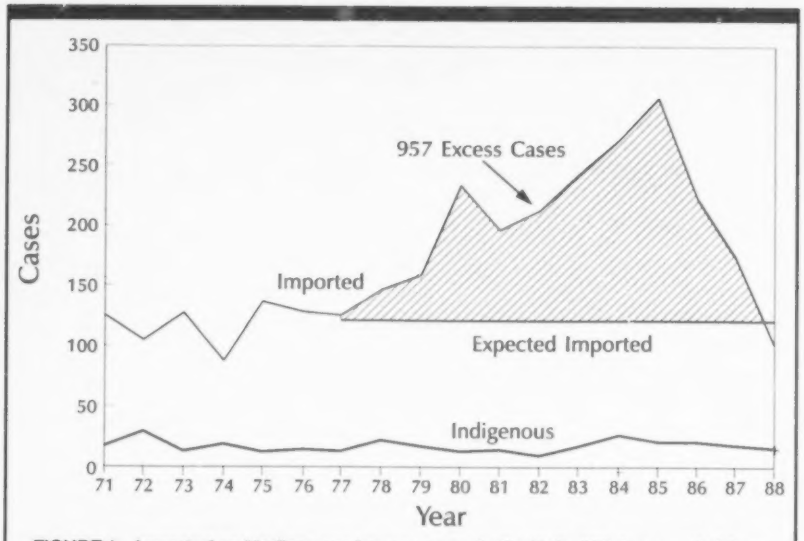


FIGURE 1—Imported and indigenous leprosy cases in the United States reported from 1971 through 1988, and expected and excess imported cases from 1978 through 1988.

West Asia—Afghanistan, Bangladesh, India, Lebanon, Pakistan, Saudi Arabia, Sri Lanka, and Yemen.

The expected annual number of imported cases for each region of the world was calculated using the number of cases reported from 1971 through 1977 as a baseline. Excess cases reported from 1978 through 1988 are defined as the number of cases observed minus the number of expected cases.

Results

Cases

From 1971 through 1988, 3554 cases of leprosy were reported in the United States. Reports of 3420 (96.2%) cases contained sufficient information to be included in the analysis; 3101 (90.7%) were imported, while 319 (9.3%) were indigenous. Of the imported cases, 2973 (95.9%) involved persons born outside the United States. The number of indigenous cases was roughly stable throughout the period, with a mean of 17.7 (range = 10 to 29) reported per year. The number of imported cases reported annually was relatively constant from 1971 through 1977 (mean = 119, range = 88 to 136), increased to 307 cases in 1985, and decreased to 102 cases in 1988 (Figure 1).

Imported Cases

Reports of imported cases were received from 47 states, Puerto Rico, and the District of Columbia. Six states reported 2468 (79.6%) of these cases: California (1306; 42.1%), Hawaii (481; 15.5%), Texas (255; 8.2%), New York (197; 6.4%), Washington (115; 3.7%), and Illinois (114; 3.7%). The mean (± SD) age of persons with leprosy was 38 (±17) years; 63.2% were male. The type of leprosy was multibacillary (lepromatous or borderline) in 67.8% of cases, 28.4% of cases were tuberculoid, and 3.8% were indeterminate.

Persons born in Southeast Asia accounted for 47.8% of all imported cases; however, using 1971 through 1977 as the baseline period to calculate the expected numbers of cases, persons from this region accounted for 73.4% of the excess cases reported from 1978 through 1988 (Table 1). Among cases from Southeast Asia, 743 (50.2%) involved persons from the Philippines and 673 (45.4%) involved persons from the three Indochinese countries of Vietnam (388 cases), Cambodia (215 cases), and Laos (70 cases). However, because of the small number of cases among Indochinese (71 cases) reported from 1971 through 1977, 69.8% of the excess cases from Southeast Asia and 51.3% of all excess cases involved persons from Indochina.

The increase in cases of leprosy imported from Indochina followed an increase in refugees from this area entering the United States (Figure 2). The first large group of refugees arrived in 1975, primarily from Vietnam,[9] followed by large groups in 1979 to 1982 consisting of refugees from Vietnam, Cambodia, and Laos.

Of the excess cases reported from 1978 through 1988, 203 (21.2%) involved persons from the Philippines. The mean annual number of Filipino immigrants to the United States from 1978 through 1988 was 45 025, compared with 34 505 from 1971 through 1977, an increase of 30.5%; the mean annual number of leprosy cases increased 61.7% from 30.0 to 48.5.

Analysis of 1-year cohorts of persons who entered the United States from 1971 through 1988 and who had onset of leprosy after immigration indicates that the highest rates of onset occurred within 12 months of arrival, with a sharp drop in onset during succeeding years of residency. For Filipino immigrants, the rate in the first year was 30.9 per 100 000 immigrants, compared with 16.8 per 100 000 refugees among Indochinese and 4.7 per 100 000 among all other immigrants.

Persons from Indochina with leprosy were younger (mean age = 33 years) than persons from the Philippines (mean age = 43 years; $P < .001$) and other areas (mean age = 38 years; $P < .001$). Onset of leprosy occurred after entry into the United States in only 53.9% of cases from Indochina, compared with 82.4% of cases from the Philippines ($P < .001$) and 78.5% of cases from all other areas ($P < .001$). Also, the mean delay from onset of illness to time of diagnosis was shorter for cases from Indochina (6 months) than for cases from the Philippines (10 months; $P < .001$) and cases from all other countries (11 months; $P < .001$). Leprosy was less commonly multibacillary (lepromatous or borderline) in persons from Indochina (50.8%) and the Philippines (64.5%) than in persons from other countries (77.0%; $P < .001$).

Indigenous Cases

Reports of indigenous cases were received from 17 states; the 4 states where leprosy is endemic reported 298 (93.4%) of these cases: Texas (191; 59.9%), Hawaii (58; 18.2%), Louisiana (40; 12.5%), and California (9; 2.8%). Of the remaining 21 indigenous cases, 10 involved persons who had resided in 1 of these 4 states. The likely place of exposure for the remaining 11 cases could not be determined from reported information. During the 18-year surveillance period, no state had a significant increase or decrease in the number of indigenous cases.

The mean (± SD) age of persons with indigenous cases of leprosy was 49 (±19) years; 53.9% were male. Race/ethnicity was reported as Hispanic for 39.5% of persons with leprosy, 35.7% were non-His-

panic White, 17.6% were Asian or Pacific Islander, and 5.3% were non-Hispanic Black (race/ethnicity was not reported for 2.8%). One hundred ninety-six (61.4%) cases were classified as lepromatous or borderline leprosy, 93 (29.2%) were tuberculoid, and 28 (8.8%) were indeterminate. The mean delay from onset of illness to diagnosis of leprosy was 10 months.

Discussion

An epidemic of imported leprosy in the United States began in 1978 and ended by 1988. The increase was primarily due to cases that occurred among the large groups of refugees that entered the United States in 1975 and 1979 through 1982 from the Indochinese countries of Vietnam, Cambodia, and Laos. The number of cases reported annually among this group decreased after the number of refugees stabilized following 1982.

Imported cases of leprosy appear to present a negligible risk for transmission in the United States; no coincident increase in indigenous leprosy occurred during this time. However, since leprosy has an incubation period of several years and since most household contacts of immigrants and refugees are likely to be immigrants and refugees themselves, it is possible that cases resulting from transmission in the United States would have been categorized as imported cases.

For all groups with imported disease, the highest rate of onset of leprosy occurred in the first year after entry into the United States, and a sharp drop was observed by the second year. Although the prevalence of leprosy in the Philippines was estimated to be 1 to 1.9 cases per 1000 population compared with 5 to 9.9 cases per 1000 population in Indochina,[1] higher rates were observed among immigrants from the Philippines than among refugees from Indochina. The increase in imported cases from the Philippines (21.2% of excess cases) was only partially explained by an increase in immigrants and may have been due, in part, to improved case reporting. However, only 17.6% of cases of leprosy among Filipinos were detected prior to entry to the United States compared with 46.1% of cases among Indochinese, perhaps reflecting a more efficient screening procedure for refugees. Also, methods of counting refugees and immigrants are different, and comparisons of rates of leprosy between these two populations should be made with caution. Refugees are eligible for special processing related to admission to the United States. A refugee's date of admission is recorded by the Immigration and Naturalization Service as the date of entry into the United States as a refugee; few enter uncounted.[10] This is in contrast to immigrants, who are recorded by the year they adjust to permanent resident status, not the year they enter the country. Not reflected in immigrant figures are persons in the United States awaiting adjustment of status and those in the country temporarily and illegally. No rate was calculated for Mexican immigrants because of an inability to accurately estimate the number of Mexicans in the United States at risk for leprosy.

Although refugees and immigrants who entered the United States from 1971 through 1988 were screened for leprosy, the majority of imported cases had onset of illness after entering the United States, with the highest rate of onset in the year following entry. Given the slow progression of clinical signs and symptoms,[4,5] it is likely that many of these persons had clinical evidence of leprosy that was not detected at the time of entry. Although most (67.8%) imported cases were lepromatous or borderline and therefore were potentially infectious, there is no evidence that imported cases of leprosy resulted in transmission in the United States. This apparent lack of transmission may be due, in part, to a combination of better housing conditions, resulting in decreased crowding, and improved nutritional status in the United States. Leprosy transmission is believed to result from prolonged and intimate contact with a person with multibacillary (lepromatous or borderline) dis-

TABLE 1—Number of Imported Cases of Leprosy in 1971 to 1988 and Excess Cases Reported in 1978 to 1988, by Region of Origin (Country of Birth)

Region	Total Cases, 1971–1988, No. (%)	Excess cases, 1978–1988, No. (%)[a]
Southeast Asia	1481 (47.8)	702 (73.4)
Mexico and Central America	720 (23.2)	126 (13.2)
Caribbean	255 (8.2)	67 (7.0)
Western Pacific	230 (7.4)	−32 (−3.3)
United States	128 (4.1)	−31 (−3.2)
South and West Asia	106 (3.4)	60 (6.3)
South America	70 (2.3)	31 (3.2)
East Asia	57 (1.8)	6 (0.6)
Africa	34 (1.1)	26 (2.7)
Europe	20 (0.6)	2 (0.2)
Total	3101	957

[a]Calculated using 1971–1977 as the baseline period.

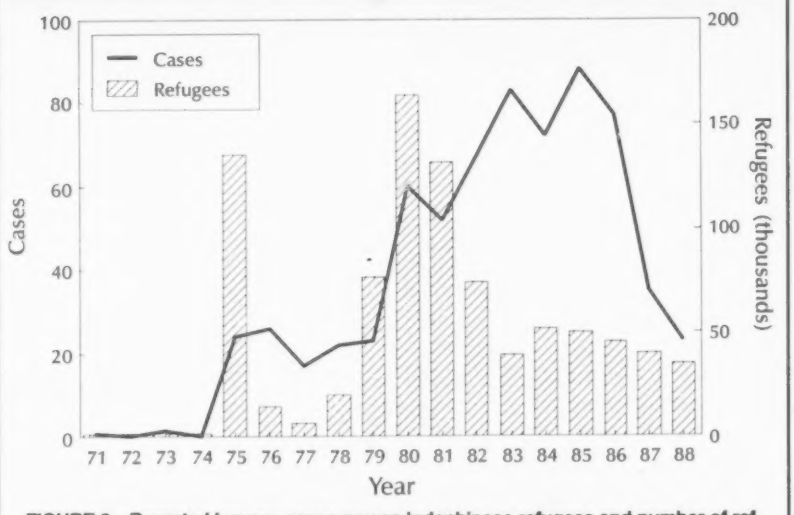


FIGURE 2—Reported leprosy cases among Indochinese refugees and number of refugees who entered the United States from 1971 through 1988.

ease; malnutrition is widely considered to predispose to leprosy.[5] In the United Kingdom, despite the diagnosis of more than 1250 cases of imported leprosy since 1951, there have been no reports of secondary cases.[11] A recent investigation of a case of imported active lepromatous leprosy in an elderly man in an English nursing home detected no secondary cases of leprosy, but immunologic studies suggested that *M. leprae* may have been transmitted (leading to subclinical infection) to 2 of 66 persons with whom the index case had daily contact.[11]

The data presented here provide additional evidence that persons with imported leprosy, including lepromatous forms, do not pose an appreciable infectious risk in the United States. Furthermore, current chemotherapy, which should be initiated promptly following diagnosis, is highly effective for all forms of leprosy; infectiousness is generally lost within 3 days of treatment with a multidrug regimen including rifampin.[3,12] Leprosy, active tuberculosis, human immunodeficiency virus infection, and five bacterial sexually transmitted diseases were classified in the Immigration and Nationality Act as "dangerous contagious diseases" for which aliens may be excluded from admission to the United States.[13,14] This study supports the proposed change to this act that would remove leprosy from the newly named list of "communicable disease of public health significance."[13,14] Continued classification of leprosy as an excludable condition unjustifiably furthers the social stigmatization historically associated with this diagnosis.

The World Health Organization has recently set as a goal the worldwide elimination of leprosy as a public health threat; elimination was defined as the reduction of prevalence to below one case per 10 000 population.[2] In the United States, continued surveillance for leprosy will be necessary to monitor changes in the distribution of indigenous cases, in particular, to further document the decrease in cases in Hawaii[15] and to determine whether the small number of reported cases in California truly represents endemic disease. Surveillance is also important to monitor trends in imported disease and to determine whether household contacts, notably children, of immigrants and refugees are at increased risk for leprosy.

Clinicians should be familiar with the clinical manifestations of leprosy and consider the diagnosis, particularly for persons from leprosy-endemic areas, both in the United States and abroad. All household contacts of persons with leprosy should be thoroughly evaluated for disease.[16] Even if indigenous leprosy is eliminated in the United States, immigrants and refugees entering the country from leprosy-endemic areas are likely to continue to present with new cases of disease. □

Acknowledgments

Portions of this paper were presented at the 30th Interscience Conference on Antimicrobial Agents and Chemotherapy, Atlanta, Ga, October 1990.

We thank Sheila F. Collin for her efforts in maintaining the leprosy surveillance system and Drs. Robert C. Hastings and Robert C. Good for valuable suggestions for the manuscript.

References

1. World Health Organization Study Group. *Epidemiology of Leprosy in Relation to Control.* Geneva: World Health Organization; 1985. WHO Technical Report Series 716.
2. World Health Organization. *UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR).* Geneva: World Health Organization; 1991. TDR news, no. 36.
3. Noordeen SK. Leprosy control through multidrug therapy (MDT). *Bull WHO.* 1991;69:263–269.
4. Meyers WM. Leprosy. In: Strickland GT, ed. *Hunter's Tropical Medicine.* Philadelphia, Pa: WB Saunders; 1991:483–494.
5. Fine P. Leprosy: the epidemiology of a slow bacterium. *Epidemiol Rev.* 1982;4: 161–188.
6. Enna CD, Jackson RR, Trautman JR, Sturdivant M. Leprosy in the United States, 1967–76. *Public Health Rep.* 1978;93:468–473.
7. Neill MA, Hightower AW, Broome CV. Leprosy in the United States 1971–1981. *J Infect Dis.* 1985;152:1064–1069.
8. Golden GS, McCormick JB, Fraser DW. Leprosy in the United States, 1971–1973. *J Infect Dis.* 1977;135:120–125.
9. Centers for Disease Control. Hansen's disease in Vietnamese refugees. *MMWR.* 1976;24:455–456.
10. US Immigration and Naturalization Service. *Statistical Yearbook of the Immigration and Naturalization Service, 1988.* Washington, DC: US Government Printing Office; 1989.
11. Dockrell HM, Eastcott H, Young S, MacFarlane A, Hussain R, Waters MFR. Possible transmission of *Mycobacterium leprae* in a group of UK leprosy contacts. *Lancet.* 1991;338:739–743.
12. Anonymous. Leprosy (Hansen's disease). In: Benenson AS, ed. *Control of Communicable Diseases in Man.* Washington, DC: American Public Health Association; 1990: 243–247.
13. Medical examination of aliens. *Federal Register.* January 23, 1991;56:2484.
14. Medical examination of aliens. *Federal Register.* May 31, 1991;56:25000.
15. Worth RM, Bomgaars MR. Immigration and leprosy in Hawaii, 1960–1981. *Int J Lepr.* 1982;50:335–341.
16. Filice GA, Fraser DW. Management of household contacts of leprosy patients. *Ann Intern Med.* 1978;88:538–542.

Public Health Briefs

Socioeconomic Status and Childhood Mortality in North Carolina

ABSTRACT

Children in families eligible to receive Aid to Families with Dependent Children (AFDC) were found to have substantially higher mortality rates than non-AFDC children. The disparity seems to be greater for Whites than for non-Whites. These results suggest the presence of social class differentials in access to and use of health care and related services and facilities, in exposure to environmental risks, and in knowledge of injury control measures. (*Am J Public Health.* 1992;82:1131–1133)

Melvin D. Nelson, Jr, MBA

Introduction

The relationship between socioeconomic status and health was first documented in this country in 1925[1] and has since been widely recognized.[2–13] The present study assesses the differences in mortality between children in families eligible to receive Aid to Families with Dependent Children (AFDC) and those not eligible for such aid from July 1, 1985, through December 31, 1988.

Methods

Lacking a more suitable indicator of poverty, the population of AFDC recipients and Medicaid participants who are eligible for AFDC was used as a surrogate. Children in the AFDC and AFDC-eligible Medicaid registries were matched to child death certificates. Death certificates for 3679 children aged 28 days through 17 years who died from July 1, 1985, to December 31, 1988, were matched with AFDC and AFDC-eligible Medicaid files for the same period. Decedents under 28 days of age were excluded because the short time between birth and death may have precluded the certification of AFDC eligibility of some otherwise potentially eligible children. Five hundred sixteen of the 3679 decedents were, at some time during this period, recipients of AFDC.

By race and age group for July 1, 1985, through December 31, 1988, unduplicated counts of AFDC eligibles aged 17 or younger were obtained. These figures were subtracted from the North Carolina population by race and age group for the same period to obtain figures for the non-AFDC population. Whether the ratio of an AFDC to a non-AFDC mortality rate was statistically different from 1.0 was determined by dividing the difference between the natural logs of the rates by the standard deviation of that difference to yield a z score.

Results

As shown in Table 1, the ratio of AFDC to non-AFDC death rates reveals that AFDC children aged 28 days through 17 years experienced a mortality rate 2.7 times that of their non-AFDC counterparts. The disparity seems to be even greater for Whites than for non-Whites.

In many cases, the cause-specific ratios are alarming. For all childhood age groups combined, AFDC children were 6.9 times more likely than non-AFDC children to die from fire and 6.1 times more likely to die from poisoning. Mortality from homicide and mortality from pneumonia or influenza were 4.7 and 4.8 times as high, respectively, for AFDC eligibles than for noneligibles. Heart disease claimed the lives of AFDC children 3.8 times more often than it claimed the lives of non-AFDC children. AFDC eligibles were 5.4 times more likely to die from perinatal conditions, 3.5 times more likely to die from congenital anomalies, 3.1 times more likely to suffer death from unintentional injuries, 2.3 times more likely to die from drowning, 1.8 times more likely to die from cancer, and 1.4 times more likely to suffer a motor-vehicle fatality.

The author is with the Center for Health and Environmental Statistics, Department of Environment, Health, and Natural Resources, Raleigh, NC.

Requests for reprints should be sent to Melvin D. Nelson, Jr., State Center for Health and Environmental Statistics, PO Box 29538, Raleigh, NC 27626-0538.

This paper was submitted to the Journal May 7, 1991, and accepted with revisions November 12, 1991.

Editor's Note. See related editorial by Davey Smith and Egger (p 1079) and commentary by Wilkinson (p 1082) in this issue.

TABLE 1—Statistically Significant Ratios of AFDC to Non-AFDC Death Rates: North Carolina, July 1, 1985 to December 31, 1988

Age Group	Cause of Death	White	Non-White	Total
28 d–1 y	Sudden infant death syndrome	1.6*	...	1.3*
	Congenital anomalies	2.5***	...	1.6*
	Perinatal conditions	3.7***	...	2.2***
	Heart disease	...	...	2.5**
	Pneumonia/influenza	...	...	2.7**
	Homicide	...	...	3.9**
	Unin non-MV injuries	3.0*	...	...
	Accid mech suffoc	6.6*	...	...
	All causes	2.1***	...	1.6***
1–4 y	Congenital anomalies	4.2***	...	2.8***
	Perinatal conditions	...	...	3.1*
	Cancer	2.8*	...	...
	Heart disease	...	...	2.7*
	Pneumonia/influenza	6.2*	...	...
	Homicide	...	...	4.1**
	Unin non-MV injuries	2.6**	3.5***	4.2***
	Fires	8.3***	4.9***	9.0***
	IICO	14.0**	...	4.7**
	Drowning	...	7.3*	...
	Motor-vehicle injuries	...	2.4**	2.6***
	All causes	2.4***	2.4***	2.9***
5–9 y	Unin non-MV injuries	3.0*	...	3.0***
	Fires	...	...	3.9**
	Drowning	...	...	3.2*
	Firearms	13.1*	...	...
	Motor-vehicle injuries	3.6**	2.7***	3.3***
	All causes	2.5***	2.0***	2.5***
10–14 y	Congenital anomalies	...	5.0*	5.2**
	Heart disease	8.2**	...	4.0*
	Pneumonia/influenza	...	...	24.2*
	Homicide	22.9***	8.4***	15.4***
	Unin non-MV injuries	...	...	2.6**
	Fires	...	...	4.5*
	Drowning	...	...	5.3***
	All causes	2.6***	2.0**	2.3***
15–17 y	Cancer	5.7***	...	2.5*
	Poisoning	57.4*	...	22.9**
28 d–17 y	Congenital anomalies	5.1***	2.8***	3.5***
	Perinatal conditions	6.7***	3.7***	5.4***
	Cancer	2.6***	...	1.8**
	Heart disease	3.1*	3.1***	3.2***
	Pneumonia/influenza	8.9***	2.2*	4.8***
	Homicide	4.3**	2.7***	4.7***
	Unin non-MV injuries	2.9***	2.4***	3.1***
	Poisoning	11.2**	...	6.1**
	Fires	6.8***	4.2*	5.9***
	IICO	7.8**	...	5.6**
	Accid mech suffoc	...	3.4*	4.3***
	Drowning	...	2.1*	2.3**
	Motor-vehicle injuries	...	1.7**	1.4**
	All causes	2.8***	2.3***	2.7***

Note. AFDC = Aid to Families with Dependent Children; Unin non-MV injuries = unintentional non-motor-vehicle injuries; Accid mech suffoc = accidental mechanical suffocation; IICO = inhalation or ingestion causing obstruction. Nonsignificant and incalculable ratios are not reported.
*$P < .05$.
**$P < .01$.
***$P < .001$.

The age-specific ratio of AFDC to non-AFDC eligibles is greatest at ages 1 through 4 (2.9) and least at 28 days to 1 year (1.6). For the 15- through 17-year-olds, the ratio of total mortality rates of AFDC to non-AFDC eligibles is not significantly different from 1.0.

For specific age groups and causes of death, some ratios of AFDC-eligible to non–AFDC-eligible death rates are remarkably high. AFDC eligibles 1 to 4 years of age are 9 times more likely to suffer death from fire than their non-AFDC counterparts. AFDC eligibles 10 to 14 years of age are 24.2 times more likely to die from pneumonia or influenza and 15.4 times more likely to be a victim of homicide than their non-AFDC counterparts. AFDC-eligible decedents 15 through 17 years of age are 22.9 times more likely to die from poisoning than their counterparts. Other age/cause-specific ratios, along with age/cause/race-specific ratios, can be found in Table 1.

Controlling for poverty, that is, comparing the mortality rates of AFDC non-Whites and AFDC Whites, the marked disparity between the races is not present as is the case in the general population. The AFDC non-White to AFDC White mortality ratios are not significantly different from 1.0 for specific age/cause groups or for all ages and causes combined.

Discussion

The overall childhood mortality rate of AFDC eligibles is nearly three times the rate of children not eligible for such aid. On an age/cause-specific basis, the disparity is even more extreme. Furthermore, no statistical difference is found between the mortality rates of non-White AFDC children and their White counterparts.

The strong association between childhood mortality and socioeconomic status points to the existence of social class differences in housing, nutrition, education, exposure to environmental risks, and access to and use of health care and related services and facilities.

That poor children die from endogenous causes such as cancer, heart disease, and pneumonia/influenza, as well as from congenital anomalies and perinatal conditions, at two to five times the rates of non-poor children suggests that parents or guardians of poor children either postpone seeking medical care until it is too late to benefit or do not have access to medical care.

The relationship of socioeconomic status to exogenous causes of death is not as well documented. That fire takes the lives of AFDC children at nearly seven times the rate of children not enrolled in AFDC suggests that substandard housing among the poor and their inability to afford smoke detectors may be a factor. Other researchers have reported a greater risk of house fires in rental property[14] and a lower rate of smoke detector installation by renters than by homeowners.[15]

AFDC-eligible children also die at a higher rate from the other exogenous causes

shown in Table 1, suggesting a greater exposure to a hazardous and violent environment. These exogenous causes also need traditional public health interventions such as legislation, education, or education.

The large differences in mortality rates between poor and nonpoor youth for these and other causes warrant further studies to determine, on a cause-specific basis, the reasons for these mortality differentials. □

References

1. Woodbury RH. *Causal Factors in Infant Mortality.* Washington, DC: US Government Printing Office; 1925. Children's Bureau Publication No. 142.
2. Amler RW, Dull HB, eds. *Closing the Gap: The Burden of Unnecessary Illness.* New York, NY: Oxford University Press; 1987: 125–129, 143–178.
3. Centerwall BS. Race, socioeconomic status, and domestic homicide, Atlanta, 1971–72. *Am J Public Health.* 1984;74:813–815.
4. Williams KR. Economic sources of homicide: reestimating the effects of poverty and inequality. *Am Sociol Rev.* 1984;49:283–289.
5. US Department of Health and Human Services. *Report of the Secretary's Task Force on Black and Minority Health, Volume I: Executive Summary.* Washington, DC: US Department of Health and Human Services; 1985.
6. Maine Department of Human Services. *Children's Deaths in Maine—1976–1980, Final Report.* Augusta, Me: Maine Department of Human Services; 1983.
7. Klerman LV. *Alive and Well? A Research and Policy Review of Health Programs for Poor Young Children.* New York, NY: National Center for Children in Poverty, Columbia University School of Public Health; 1991.
8. Antonovsky A. Social class, life expectancy, and overall mortality. *Milbank Q.* 1967;45:31–73.
9. Cassel J. The contribution of the social environment to host resistance. *Am J Epidemiol.* 1976;104:107–123.
10. Pratt L. The relationship of socioeconomic status to health. *Am J Public Health.* 1971; 61:281–291.
11. Syme SL, Berkman LF. Social class, susceptibility and sickness. *Am J Epidemiol.* 1976;104:1–8.
12. Stockwell EG, Swanson DA, Wicks JW. Economic status differences in infant mortality by cause of death. *Public Health Rep.* 1988;3:135–142.
13. Mare RD. Socioeconomic effects on child mortality in the United States. *Am J Public Health.* 1982;72:539–547.
14. Ducik S, Ghezzo HR. Epidemiology of accidental home fires in Montreal. *Accident Anal Prev.* 1980;12:66–73.
15. Miller RE, Reisinger KS, Blatter MM, et al. Pediatric counseling and subsequent use of smoke detectors. *Am J Public Health.* 1982;72:392–393.

Does Equal Socioeconomic Status in Black and White Men Mean Equal Risk of Mortality?

Julian E. Keil, MS, DrPH, Susan E. Sutherland, MS, Rebecca G. Knapp, PhD, and Herman A. Tyroler, MD

ABSTRACT

Although concerns have been expressed that mortality from coronary disease and all other causes is greater among Blacks than Whites, we hypothesized that, when socioeconomic status is adequately considered, mortality inequalities between Blacks and Whites are insignificant. The study population was a random sampling of Black and White men who were 35 years of age or older when recruited into the Charleston Heart Study in 1960. Education level and occupational status at baseline were used to compare mortality over the ensuing 28 years between Black and White men, who were classified as low or high socioeconomic status. In no instance were Black-White differences in all-cause or coronary disease mortality rates significantly different when socioeconomic status was controlled. We conclude that socioeconomic status is an important predictor of mortality and that, when socioeconomic status is considered, differences in Black-White mortality rates may be small. (*Am J Public Health.* 1992;82:1133–1136)

Introduction

In the early 1970s Terris[1] expressed disdain for evaluation of health statistics by race, contending that the appropriate variable for comparison was socioeconomic status. Despite evidence by Antonosky, Kitagawa, Kaplan, Feldman, Marmot, and Haan[2–7] of an inverse relationship between socioeconomic status and mortality, studies have continued to focus on ethnicity. Such studies[8–9] identify a racial difference in risk factors for mortality but, in most instances, use the concept of socioeconomic status to explain the reported difference. Nevertheless, these studies have served social and humanitarian purposes by identifying underserved populations at particular risk of disease or death. In the Charleston Heart Study cohort, low socioeconomic status was found to be a significant predictor of the incidence of hypertension, while skin color was not.[10] Recently, Gillum[11] has shown that mortality from coronary heart disease or from all other causes among Blacks has been greater than among Whites. The purpose of this report is to provide additional evidence from the Charleston Heart Study to support the hypothesis that socioeconomic status is a key predictor of mortality when ethnicity is controlled.

Materials and Methods

The study population in Charleston, SC, was a random sampling of Black and White men who were 35 to 74 years of age

Julian E. Keil, Susan E. Sutherland, and Rebecca G. Knapp are with the Department of Biostatistics, Epidemiology, and Systems Science, Medical University of South Carolina, Charleston. Herman A. Tyroler is with the Department of Epidemiology, School of Public Health, University of North Carolina, Chapel Hill, NC.

Requests for reprints should be sent to Julian E. Keil, MS, DrPH, Charleston Heart Study, Room 908 Harborview Towers, Medical University of South Carolina, Charleston, SC 29425-2239.

This paper was submitted to the Journal July 3, 1991, and accepted with revisions December 31, 1991.

Editor's Note. See related editorial by Davey Smith and Egger (p 1079) and commentary by Wilkinson (p 1082) in this issue.

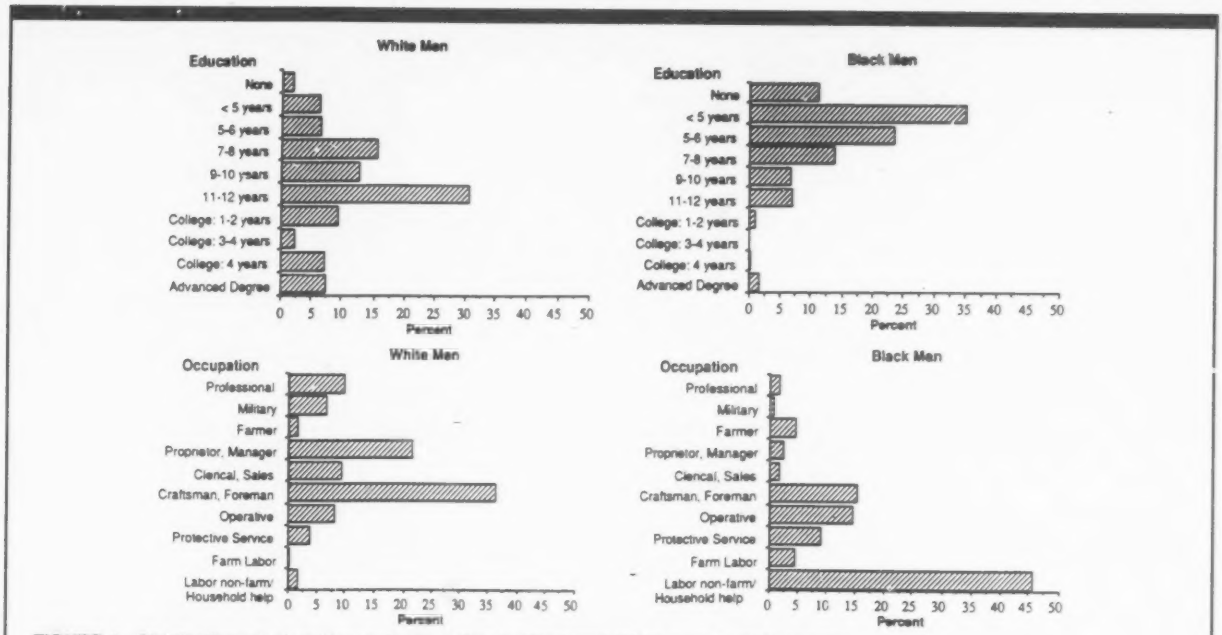


FIGURE 1—Distribution of education and occupation in Black and White men in 1960 sample.

TABLE 1—Status of the Cohort in 1988

	1960 Cohort	Deceased							
		Coronary Heart Disease[a]		Other Causes		Alive		Unable to Locate	
	n	n	%	n	%	n	%	n	%
White men	653	131	20	240	37	267	41	15	2
Black men	333	50	15	157	47	123	37	3	1
High socioeconomic status Black men	102	8	8	26	25	68	67	0	0

[a]Codes 410–414 of the *International Classification of Diseases*, 9th version, from nosologist coding of underlying cause of death.

in 1960. Details of the baseline examination and sampling plan have been previously published.[12] In 1963, the cohort was supplemented with 102 Black men of high socioeconomic status who had been recruited by peer nomination to provide an extended range of economic status in Blacks. Between 1960 and 1988, there were four recalls of the cohort. Table 1 provides the vital status of the cohort through 1988.

While there was a wide range of education and occupations among White men, more than 85% of randomly selected Black men had 8 years or less of education and 75% were in the lower occupation groups of laborer, protective service worker, craftsman, or operative (Figure 1). To assess the influence of both race and socioeconomic status, comparisons of age-adjusted coronary heart disease and all-cause mortality rates[13] were made in four groups: Black, White, low socioeconomic status, and high socioeconomic status. Low socioeconomic status was defined as having 0 to 8 years of education and being in the lower occupation group. High socioeconomic status was defined as having some college education and an occupation of proprietor or professional.

The participants of low socioeconomic status included in the analyses were from the random sampling of 1960. For high socioeconomic status, the White men were from the random sampling of 1960, while most of the Black men were from the special cohort recruited in 1963. Baseline characteristics are described in Table 2.

Results

Table 3 provides a comparison of mortality among Black and White male participants in the Charleston Heart Study when socioeconomic status criteria are strictly met, and also provides estimates of the effect of socioeconomic status while controlling for race. White men had higher rates than Black men, regardless of socioeconomic status or cause of death. In no instance was the White–Black difference in rates statistically significant, although the White–Black ratio for coronary heart disease mortality among men of low socioeconomic status was 1.8 (7.1/4.0). Among Whites for all causes of death, men of low socioeconomic status had rates 1.8 times those of men of high socioeconomic status. The rate ratio was 2.1 for coronary heart disease mortality. Among Blacks, the low–high socioeconomic status ratio was 2.1 for all-cause mortality and 1.6 for coronary heart disease mortality.

Discussion

There is evidence in both White and Black men that socioeconomic status has a powerful influence on coronary mortality and death from all causes. In every comparison controlling for socioeconomic status, the rates for Black men were slightly, but not significantly, lower than for White men.

TABLE 2—Charleston Heart Study Baseline Characteristics of Men Aged 35–74 by Socioeconomic Status

	Low Socioeconomic Status[a]						High Socioeconomic Status[b]					
	White (n = 151)			Black (n = 258)			White (n = 115)			Black (n = 77)		
	Mean	SD	%	Mean	SD	%	Mean	SD	%	Mean	SD	%
Age	51.5	9.34		49.1	9.75		47.9	9.83		47.4	9.63	
Systolic blood pressure	143.6	23.58		153.6	30.09		139.6	20.25		139.9	18.85	
Diastolic blood pressure	85.3	10.88		91.5	14.18		83.8	9.40		85.6	12.76	
Cholesterol	229.9	41.98		222.8	51.76		242.1	41.24		251.4	45.15	
Body mass index	25.6	4.23		25.3	4.32		25.1	2.91		27.2	4.72	
Education, y	6.0	1.9		4.2	2.6		14.9	2.6		15.2	2.6	
Smoking			65.8			56.2			56.3			51.9
Diabetes			2.0			4.3			2.6			2.6
Electrocardiogram LVH			2.6			9.7			0.9			2.5
Medical care												
Private physician			87.8			65.1			98.3			98.7
Clinic			3.4			14.9			1.8			0.0
None			8.8			20.0			0.0			1.3
Occupation												
Laborer			4.6			60.6			...			...
Other			95.4			39.4			...			...
Professional			...			...			51.3			71.4
Proprietor			...			...			48.7			28.6

[a]Education = less than 8 years; occupation = craftsman, operative, protective, farm laborer, or other laborer.
[b]Education = some college or greater; occupation = professional or proprietor.

Black men of low socioeconomic status had lower baseline levels of education, higher blood pressure, more electrocardiogram left ventricular hypertrophy, and less access to medical care, but had lower cholesterol and were younger, than the White men. Age adjusting the risk factors of the low socioeconomic status group had the effect of increasing the risk factors of Black men, yet their mortality rates were lower. White and Black men of high socioeconomic status had comparable characteristics. Blood pressure was higher and cholesterol was lower in the low than in the high socioeconomic status ethnic peer group. In both socioeconomic status groups, Black men had lower mortality rates than White men. Black men of lower socioeconomic status had 13.7% lower all-cause rates and 43.1% lower rates for coronary heart disease mortality than Whites. Black men of high socioeconomic status had 24% lower all-cause and 32% lower coronary heart disease mortality than White men of high socioeconomic status.

TABLE 3—Mortality Rates,[a] 1960–1988: White and Black Men

	All-Cause Mortality			Coronary Heart Disease Mortality		
	n	Age-Adjusted Rate	(95% CI)	n	Age-Adjusted Rate	(95% CI)
Low socioeconomic status						
White men (n = 151)	101	20.9	(13.5, 28.3)	33	7.1	(2.2, 11.9)
Black men (n = 258)	159	18.0	(14.3, 21.7)	36	4.0	(2.2, 5.9)
High socioeconomic status						
White men (n = 115)	59	11.5	(7.5, 15.4)	22	3.5	(2.0, 5.0)
Black men (n = 77)	30	8.7	(5.6, 11.8)	8	2.4	(0.7, 4.1)

[a]Age-adjusted rate per 1000 person years.

Our findings may be mediated by the Black–White mortality crossover effect reported by Wing[14] that suggests that once Blacks attain age 65 they have preferred survival status. It has also been postulated that socioeconomic status may have opposite effects during the earlier and later periods of the coronary heart disease epidemic. This concept, suggestive of a socioeconomic status crossover, was demonstrated by Marmot[15,16] to show "that as the CHD [coronary heart disease] epidemic grew in industrialized countries, it affected first the more affluent classes and then percolated through the social classes." A similar finding in the Evans County, Georgia, cohort was reported by Morgenstern.[17] For the Charleston Heart Study cohort, a socioeconomic status crossover effect was not detected. However, since Black men of low socioeconomic status had lower rates and fewer years of education and were more likely to be laborers than White men, a protective effect for these men is suggested. Our findings of lower mortality in Black men of low socioeconomic status may suggest that they were acculturating at a rate somewhat slower than White men of low socioeconomic status, providing evidence for Marmot's thesis.

Finally, more positive interpretations of our results are difficult because of the relatively small sample size and the commensurate large confidence intervals around the estimates of mortality.

While our results are comforting in that Black men per se may not have higher mortality than White men, they may also provide further evidence of the validity of

Antonosky's exclamation[2] of the 1960s: "The inescapable conclusion is that class influences one's chance of staying alive." □

Acknowledgments

This research was supported by National Heart, Lung, and Blood Institute Grant 5 R01 HL31397-08.

The authors acknowledge the contribution of the late Dr. Edwin Boyle, Jr., founder of the Charleston Heart Study, to this work.

References

1. Terris M. Desegregating health statistics. *Am J Public Health.* 1973;63:477–480.
2. Antonosky A. Social class, life expectancy and overall mortality. *Milbank Q.* 1967;45: 31–73.
3. Kitagawa EM, Hauser PM. *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology.* Cambridge, Mass; Harvard University Press; 1973.
4. Kaplan GA, Haan MN, Syme SL, et al. Socioeconomic status and health. *Am J Prev Med.* 1987;3(suppl):125–129.
5. Feldman JJ, Makuc DM, Kleinman JC, et al. National trends in educational differentials in mortality. *Am J Epidemiol.* 1989; 129:919–933.
6. Marmot MG, Kogevinas M, Elston MA. Social/economic status and disease. *Annu Rev Public Health.* 1987;8:111–135.
7. Haan MN, Kaplan GA, Syme SL. Socioeconomic status and health. Old observations and new thoughts. In: Bunker JP, Gumby PS, Kehrri BH, eds. *Pathways to Health. The Role of Social Factors.* Menlo Park, Calif: Henry J. Kaiser Foundation; 1989.
8. Klag MJ, Whelton PK, Coresh J, et al. The association of skin color with blood pressure in US blacks with low socioeconomic status. *JAMA.* 1991;265:599–602.
9. Dressler WW. Social class, skin color and arterial blood pressure in two societies. *Ethnicity & Disease.* 1991;1(1):60–77.
10. Keil JE, Tyroler HA, Boyle E. Hypertension: effects of social class and racial admixture. The results of a cohort study in the Black population of Charleston, SC. *Am J Public Health.* 1977;67:634–639.
11. Gillum RF. Cardiovascular disease in the United States. An epidemiologic overview. In: Saunders E, Brest AN, eds. *Cardiovascular Diseases in Blacks.* Philadelphia, Pa: FA Davis Co; 1991.
12. Boyle E. Biological patterns in hypertension by race, sex, body weight and skin color. *JAMA.* 1970;213:1637–1643.
13. Bergstralh EJ, Offord KP, Kosanke JL, Augustine GA. PERSONYRS: A SAS procedure for person years analysis. In: *Proceedings of the 11th Annual SAS Users Group International Conference, Atlanta, GA, February 9–12, 1986.* Cary, NC: SAS Institute Inc; 1986.
14. Wing S, Manton KG, Stallard E, et al. The Black/White mortality crossover: investigation in a community-based study. *J Gerontol.* 1985;40:78–84.
15. Marmot M. Socioeconomic determinants of CHD mortality. *Int J Epidemiol.* 1989; 18(suppl 1):196–202.
16. Marmot M, Adelstein AM, Robinson N, Rose GA. Changing social-class distribution of heart disease. *Br Med J.* 1978;2: 1109–1112.
17. Morgenstern H. The changing association between social status and coronary heart disease in a rural population. *Soc Sci Med.* 1980;14A:191–201.

Job Strain and Mortality in Elderly Men: Social Network, Support, and Influence as Buffers

Anders Falk, MD, Bertil S. Hanson, MD, Sven-Olof Isacsson, MD, and Per-Olof Östergren, MD

ABSTRACT

The aim was to investigate whether job strain affects mortality in a representative population of elderly men, and whether social network and social support outside the workplace can buffer the negative health effects of job strain. A higher relative mortality risk (RR) was found among men exposed to job strain (RR = 1.7). The combination of exposure to job strain and seven different measures of weak social network and social support was associated with a further increased RR ranging from 2.1 to 4.6. (*Am J Public Health.* 1992;82:1136–1139)

Introduction

In 1979, Karasek proposed a two-dimensional model for measuring psychosocial work conditions, in which the combination of high psychological job demands and low personal schedule freedom (job strain) was found to be associated with depression, exhaustion, and job dissatisfaction.[1] After that, several studies found associations between job strain and cardiovascular disease.[2–6] Job support was later added as a third concept in a three-dimensional model (iso-strain).[7–8]

Social network and social support have been shown to affect morbidity and mortality,[9–13] and it has been suggested that they may have a buffering function on the negative health effects of different stressors.[14]

Previous studies on the health effects of job strain have all been performed on active, working populations. No studies have investigated the chronicity of job strain and its effects on health in elderly people.

The aims of this study are therefore to investigate whether exposure to job strain during the active working period affects mortality in elderly men after retirement and to investigate if the social network and social support outside the place of work can buffer the negative health effects of job strain.

The authors are with the Department of Community Health Sciences of Lund University, Malmö General Hospital, in Sweden.

Requests for reprints should be sent to Dr. Bertil S. Hanson, Department of Community Health Sciences, Lund University, Malmö General Hospital, S-214 01 MALMÖ, Sweden.

This paper was submitted to the Journal April 9, 1991, and accepted with revisions December 31, 1991.

Methods

All 621 men born in even months in 1914 and residing in the city of Malmö, Sweden (about 230 000 inhabitants), during 1982 to 1983 formed the study sample. Five hundred men (80.5%) agreed to participate, 23 only partially.[15]

The investigation took place during 1982 to 1983, 3 years after the participants' retirement. Interviews, questionnaires, and physical and physiological examinations were used to obtain information about health status, life-style, the psychosocial working environment, social network, and social support.

The dichotomous job strain variable was constructed by combining high job demands and low personal schedule freedom according to a method described by Karasek et al.[3] Affirmative answers to two questions (Was your job hectic? Was your job psychologically demanding?) were used to define high job demands. Low personal schedule freedom was defined as negative answers to at least two of three questions (Could you make at least one private phone call of about 10 minutes during working hours? Could you go on a private errand for about 30 minutes without asking a superior? Could you receive a private visitor for about 10 minutes during working hours?).

The indices measuring different aspects of social network and social support have been described in detail previously (see appendix).[16-17] Social network includes social anchorage (sense of anchorage within formal and informal groups), contact frequency (contacts with family, neighbors, friends), and social participation (degree of participaion in social activities).

Social support includes emotional support (opportunities to receive care and personal encouragement), informational support (access to guidance and information), and material support (access to practical services and material resources).

Low social network and social support was defined as the lowest third of the distributions of the indices.[17]

Mortality of the cohort was followed continuously until February 28, 1989, using the Cause of Death Register both at the Department of Community Health Sciences in Malmö and at the Swedish National Central Bureau of Statistics.

TABLE 1—Relative Mortality Risks between High and Low Categories of Social Network and Social Support

	Relative Mortality Risks	95% CI
Social network		
Contact frequency	1.2	0.8, 1.8
Social anchorage	1.6	1.1, 2.3
Social participation	1.2	0.8, 1.8
Adequacy of social participation	1.9	1.1, 3.4
Social support		
Availability of material and informational support (low/high)	1.4	0.9, 2.0
Availability of emotional support	1.8	1.2, 2.8
Adequacy of emotional support	1.3	0.8, 2.0

Note. Figures are for men born in 1914 and residing in Malmö, Sweden, during 1982 to 1983.

The relative mortality risks (RR) were calculated using risk odds ratios and multivariately using logistic regression analysis. Homogeneity of the RRs was tested with a chi-square test for homogeneity.

Results

Eighty-seven men (18.2%) of the 477 full participants died during the follow-up. Men exposed to high demands had an RR of 1.3 (0.8–2.0), and those exposed to low personal schedule freedom had an RR of 0.7 (0.5–1.2). Men exposed to job strain showed a higher all-cause mortality rate (23/91) than those not exposed (64/386) (RR = 1.7 [1.0–2.3]). Adjustment for social class, cardiovascular risk factors, and

TABLE 2—Relative Mortality Risks (RR) in Groups with Different Combinations of Job Strain and Social Network and Social Support Characteristics

	Social Network								Social Support					
	Contact Frequency		Social Anchorage		Social Participation		Adequacy of Social Participation		Availability of Material and Informational Support		Availability of Emotional Support		Adequacy of Emotional Support	
	RR	CI[a]	RR	CI	RR	CI	RR	CI	RR	CI	RR	CI	RR	CI
High social network, social support, and														
no job strain	1.0	...	1.0	...	1.0	...	1.0	...	1.0	...	1.0	...	1.0	...
job strain	1.4	0.7, 2.7	1.0	0.5, 2.1	1.6	0.8, 3.4	1.3	0.6, 2.6	1.3	0.6, 2.9	1.8	1.0, 3.4	1.2	0.6, 2.4
Low social network, social support, and														
no job strain	1.1	0.6, 2.0	1.2	0.7, 2.3	1.3	0.7, 2.2	1.7	1.0, 3.1	1.3	0.7, 2.2	2.3	1.2, 4.4	1.0	0.5, 2.0
job strain	2.9	1.2, 6.7	4.6	2.1, 10.0	2.1	1.0, 4.6	4.4	2.0, 9.5	2.6	1.3, 5.3	2.7	1.1, 6.9	3.6	1.6, 8.2
p^b	.09		.002		.23		.002		.07		.01		0.02	

Note. Figures are for men born in 1914 and residing in Malmö, Sweden, during 1982 to 1983.
[a]CI = 95% confidence intervals.
[b]Chi-square test for homogeneity.

variations in follow-up time did not change the relative risk very much (RR = 1.6 [1.0–2.5]).

The RRs for men with a low score on the social network and social support indices were compared with those for men with a high score (Table 1). All indices showed RRs higher than 1.0.

A multicategorial analysis was performed to investigate if there was evidence of a buffering effect (i.e., effect modification) between job strain and social network and social support (Table 2). The RRs regarding exposure to job strain were highest in the groups exposed to low social network and social support, ranging from 2.1 to 4.6, using the group with no job strain/high social network and support as reference (RR = 1).

Discussion

Considering the high participation rate together with the fact that the health differences between the participants and the nonparticipants were rather small, the participants could be regarded as fairly representative of this population of elderly men.[15]

The individual's view of his personal schedule freedom probably was not affected too much by the fact that 3 years or more had passed between being in active work and participating in this study. The assessment of job demand, however, was more indirect and measured the sum of several components of the psychosocial work conditions. A nondifferential misclassification because of a fading recollection of the actual job situation, however, meant an underestimation of the true mortality risks.

The association between job strain and mortality could be the result of social class differences or of a higher prevalence of risk factors in the strain group. However, such effects did not show in the multivariate analysis, consistent with other studies.[4]

The health effect of social network and social support may be mediated through different causal pathways. One such pathway is described by the buffering hypothesis; i.e., social network and social support are mobilized in response to environmental stressors.[14] Men with a combination of high job strain and insufficient social network and social support showed the highest mortality risks, thus supporting a buffering effect.

Our findings indicate that exposure to job strain also influences the risk of mortality after retirement, and that social network and social support outside work seem to buffer the negative health effects of job strain. □

Acknowledgments

This study has been supported by a grant from the Swedish Medical Research Council and the medical faculty at Lund University.

References

1. Karasek R. Job demands, job decision latitude, and mental strain: implications for job redesign. *Admin Sci Q.* 1979;24:285–307.
2. Alfredsson L, Karasek R, Theorell T. Myocardial infarction risk and psychosocial environment: an analysis of the male Swedish working force. *Soc Sci Med.* 1982; 3:463–467.
3. Karasek R, Baker D, Marxer F, Ahlbom A, Theorell T. Job decision latitude, job

APPENDIX

The following items were included in the interview questionnaire that measured the different aspects of social network and social support.

Social Anchorage

Would you say that you are rooted and have a feeling of familiarity with your neighborhood?

How many years have you lived in your neighborhood?

If you were able to move now, would you do so?

Do you belong to a group of friends that does things together, i.e., play cards, listen to music, go on picnics, etc.?

In your daily life, do you have use for the knowledge and skills you have acquired during your life?

Do you feel that you are of great importance to other people and that you have a useful function in society?

Are you a member of any organizations or clubs?

Are you in any position of trust in any organizations or clubs?

Contact Frequency

How often do you meet any of these persons (children, kin, neighbors, friends, close friends, workmates) at your or at their homes?

Social Participation

How often in the last year have you taken part in study circles and other courses?

How often do you take part in meetings of different organizations, clubs, etc.?

Have you done any of the following during the last 12 months: written an article, taken an active part in the discussion at a meeting, taken part in a political meeting, tried to influence a decision regarding a political question, appealed against a decision, contacted a consumers' association, gone to a salesman or shop in order to complain, taken part in a political or union rally?

How often in the last year have you been to parties at your friends', relatives', or neighbors' homes?

How often in the last year have you been to church?

Did you vote in the referendum concerning nuclear power?

Adequacy of Social Participation

Are you satisfied with your opportunities for participating in different activities?

How important do you think that your activities are to you?

How often do you find it a problem to occupy yourself?

Availability of Material and Informational Support

Is there anyone in your neighborhood from whom you can borrow things or exchange services?

Is there anyone in your neighborhood from whom you can get help if you fall ill?

If you need help with something for 24 hours, do you have anyone (other than wife) you can ask?

Do you know anyone who can help you to write an official letter or appeal against a decision made by some authority?

Do you know where to go to get help in writing an official letter or appealing against a decision made by some authority?

Availability of Emotional Support

Do you have any friends or relatives who you like very much and who like you very much?

Do you have any really close friends with whom you feel intimate and with whom you can discuss anything?

If you have continued to work, is it because you want to feel that you are a valuable and important person?

When you have personal problems of any kind, do you have any close friend or relative to whom you can turn to discuss your problems?

Adequacy of Emotional Support

How often do you feel lonely?

Do you have the feeling that people appreciate what you do?

Do you have enough good friends to be with?

Do you think that you see your children too often or too rarely?

demands and cardiovascular disease: a prospective study of Swedish men. *Am J Public Health.* 1981;71:694–705.
4. Alfredsson L, Theorell T. Job characteristics of occupations and myocardial infarction risk-effect of possible confounding factors. *Soc Sci Med.* 1983;17:1497–1503.
5. Alfredsson L, Spetz C, Theorell T. Type of occupation and near future hospitalization for myocardial infarction and some other diagnoses. *Int J Epidemiol.* 1985;14:378–388.
6. Karasek R, Theorell T, Schwartz J, Schnall P, Pieper C, Michela J. Job characteristics in relation to the prevalence of myocardial infarction in the US Health Examination Survey (HES) and the Health and Nutrition Examination Survey (HANES). *Am J Public Health.* 1988;78:910–918.
7. Johnson J. *The Impact of Place of Work Social Support, Job Demands and Work Control upon Cardiovascular Disease in Sweden,* Report Number 1. Stockholm, Sweden: University of Stockholm; 1986. Doctoral dissertation.
8. Åstrand N-E. *Prediction of Sick Leave, Early Retirement, Unemployment, Labour Turnover and Mortality in a Swedish Male Working Population: A Prospective Study.* Malmö, Sweden: Lund University; 1988. Doctoral dissertation.
9. Berkman LF, Syme SL. Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. *Am J Epidemiol.* 1979;109:186–204.
10. Seeman TE, Kaplan GA, Knudsen L, Cohen R, Guralnik J. Social network ties and mortality among the elderly in the Alameda County study. *Am J Epidemiol.* 1987;126:714–723.
11. Schoenbach VJ, Kaplan BH, Fredman L, Kleinbaum DG. Social ties and mortality in Evans County, Georgia. *Am J Epidemiol.* 1986;123:577–591.
12. Kaplan GA, Salonen JT, Cohen RD, Brand RJ, Syme SL, Puska P. Social connections and mortality from all causes and cardiovascular disease: prospective evidence from eastern Finland. *Am J Epidemiol.* 1988;128:370–380.
13. Hanson BS, Isacsson S-O, Janzon L, Lindell S-E. Social network and social support influence mortality in elderly men: the prospective study of "men born in 1914," Malmö, Sweden. *Am J Epidemiol.* 1989;130:100–111.
14. LaRocco JM, House JS, French JRP Jr. Social support, occupational stress and health. *J Health Soc Behav.* 1980;21:202–218.
15. Janzon L, Hanson BS, Isacsson S-O, Lindell S-E, Steen B. Factors influencing participation in health surveys: results from the prospective population study "Men born in 1914" in Malmö, Sweden. *J Epidemiol Community Health.* 1986;40:174–177.
16. Hanson BS, Östergren P-O. Different social network and social support characteristics, nervous problems and insomnia: theoretical and methodological aspects on some results from the population study "men born in 1914," Malmö, Sweden. *Soc Sci Med.* 1987;25:849–859.
17. Hanson BS. *Social Network, Social Support and Health in Elderly Men: A Population Study.* Malmö, Sweden: Lund University; 1988. Doctoral dissertation.

Mortality and the Attribution of Health Problems to Aging among Older Adults

William Rakowski, PhD, and Tom Hickey, DrPH

ABSTRACT

This study examined the association between mortality and attributing health problems to aging among 1391 respondents from the Longitudinal Study of Aging who indicated difficulty with activities of daily living. Of this number, 72 persons attributed impairment primarily to "old age." Logistic regression controlling for demographics, physical health problems, self-rated health, and social involvements showed an association with mortality (adjusted odds ratio = 1.78, CI = 1.05, 3.00). Attributing health problems to aging may carry a risk of adverse health events. (*Am J Public Health.* 1992;82:1139–1141)

Introduction

Research in gerontology continues to try to distinguish effects of disease from effects of aging on physical health and functioning.[1–3] Public information and professional education increasingly emphasize that attributing health problems to supposedly inevitable consequences of old age is inappropriate. The current emphasis on disease prevention and health promotion with older persons[4–6] reflects a growing recognition that chronological aging per se imposes significantly fewer limitations than was previously assumed.

It is important to understand the potential consequences of laypersons, professionals, and the general public attributing health problems to aging. Leventhal and Prohaska[7] report that attribution to aging is associated with a tendency to delay seeking care. Gjorup et al.[8] found that attribution to aging exceeds 60% for some conditions and is associated with delay in contacting a physician, while Brody and Kleban[9] note that such attribution is associated with reluctance to discuss health problems with other people or health care providers. It was therefore expected in this study that persons who made attributions to aging would have higher mortality.

Methods

Sample

The sample for this investigation was self-respondents at baseline in the 1984–1988 Longitudinal Study of Aging (LSOA) who reported difficulty with one or more activities of daily living (ADLs) (n = 1391). The baseline for the LSOA was the Supplement on Aging to the 1984 National Health Interview Survey. All LSOA participants

William Rakowski is with the Center for Gerontology and Health Care Research and the Department of Community Health, Brown University, in Providence, RI. Tom Hickey is with the Department of Health Behavior and Health Education, School of Public Health, University of Michigan, in Ann Arbor.

Requests for reprints should be sent to Dr. William Rakowski, Box G-B221, Center for Gerontology and Health Care Research, Brown University, Providence, RI 02912.

This paper was submitted to the Journal July 1, 1991, and accepted with revisions October 1, 1991.

TABLE 1—Bivariate Association between Mortality and Attribution of ADL Impairment to Aging

Independent Variable	Number	Percent Died	Odds Ratio	95% CI
ADL attributions to aging				
None	1319	30.0	(reference category)	
One or more	72	45.8	1.97	1.22, 3.18

TABLE 2—Multivariate Association between Mortality and ADL Attribution to Aging: Results from Logistic Analysis

Independent Variable	Adjusted Odds Ratio (AOR)	95% CI
ADL attributions to aging		
None	(reference category)	
One or more	1.78	1.05, 3.00

$-2 \times$ log-likelihood ratio change = 4.58, $df = 1$, $P < .05$

Note. Other predictors with 95% CIs excluding 1.00 are cancer (AOR = 1.54, 1.05–2.15), diabetes (AOR = 1.74; 1.23–2.44), male sex (AOR = 1.77; 1.33–2.36), age 80 to 84 (AOR = 1.86; 1.32–2.63), age 85+ (AOR = 2.32; 1.56–3.44), no. instrumental ADLS (AOR = 1.20; 1.10–1.30), self-rated health (AOR = 1.21; 1.07–1.36), body mass index (AOR = 1.18; 1.08–1.29), and social involvement (AOR = 1.31; 1.09–1.59).

were aged 70 and older in 1984. The 1984 interviews were conducted by Census Bureau staff at the individual's residence; the 1988 follow-up was conducted primarily by computer-assisted telephone interview.[10–12]

Independent Variable: Attribution of Problems to Aging

The ADLs assessed by the LSOA comprise dressing, bathing, using the toilet, eating, walking, getting outside, and transferring, and indicate basic functional health status.[13,14] A report of ADL difficulty was followed by questions about degree of difficulty, receipt of help, use of aids, and perceived cause(s) of an impairment. The attribution question allowed giving multiple causes for the impairment; if multiple causes were mentioned, respondents were asked to select one as the main cause. Therefore, persons classified for this report as attributing a problem to aging were those who chose aging as the main cause. The independent variable was defined as no ADLs attributed to aging versus one or more ADLs attributed to aging. (Note: Attribution to aging was not requested for physical health problems such as stroke, heart disease, cancer, etc.)

Dependent Variable: Vital Status

Mortality was based on the National Center for Health Statistics' "best estimate" of vital status, derived from tracking information through December 1988. In the analysis sample of 1391 persons, there were 429 deaths (30.8%).

Covariates for Logistic Analysis

Several other variables from the LSOA were used as covariates in logistic regression analysis. They included any heart problem, any circulatory problem, diabetes, stroke, cancer, gender-specific body mass index, number of limitations to instrumental ADLs, age, gender, education, race, living alone or with others, self-rated health, and an indicator for social contacts (based on group or social events out-of-home, volunteer work, and attendance at religious services).

Results

Bivariate Associations

Table 1 presents the bivariate associations between mortality and the attribution of ADL problems to old age over the 1984 to 1988 period. The number of persons who cited only old age as a cause was small (n = 72). However, an elevated risk of mortality did appear to exist.

Multivariate Associations

Results of logistic regression are shown in Table 2. Analyses were done with BMDP and unweighted data. Even controlling for the covariates listed above, the attribution of ADL difficulty to aging was associated with a higher risk of mortality, as denoted by the adjusted odds ratio (OR) and change in the log-likelihood ratio due to attribution. Other covariates associated with higher mortality were having had cancer or diabetes, being age 80 or over, being male, having difficulty with more instrumental ADLs, being in the lowest quintile for body mass index, having less favorable self-rated health, and having fewer social involvements.

Correlates of Attribution to Aging

In other analyses, few variables from the covariates listed above and from other questions from section T ("Health Opinions") of the baseline survey were found to be correlates of attributing ADL difficulty to aging. Table 3 presents bivariate associations for those variables. Several associations suggested greater aging attribution when health ratings were favorable. Multiple logistic analyses using none versus one or more attributions as the dependent variable indicated a greater likelihood of attribution for persons aged 80 to 84 (adjusted odds ratio [AOR] = 2.65; 95% CI = 1.16, 6.06) and aged 85 or over (AOR = 6.57; 95% CI = 2.95, 14.70), relative to persons aged 70 to 74 at baseline.

Discussion

Results of this investigation suggest a potential adverse consequence of attributing at least some types of health problems to old age as opposed to a specific disease. As hypothesized by Kart,[15(p78)] "Overattribution of symptoms to the aging process directs the attention of the elderly person away from real disease and/or environmental factors that may affect health. Such misattributions may have tragic consequences." Although the number of persons making such attributions was small, the database was a national sample and results are consistent with the above hypothesis and with data from earlier studies.[7–9]

In actuality, the number of persons making attributions to aging was probably underestimated because respondents were asked to choose a main cause if several were given. The group identified here is therefore selective and may represent individuals for whom attribution to old age has special significance. Their attribution may denote a cumulative judgment of "how bad things really are." Attributions to aging may be one factor differentiating "normal" from "successful" aging, as discussed by Rowe and Kahn,[1] and may also be another manifestation of self-rated health.[16] Global self-rated health, along

with social involvement, predicted mortality even in this impaired sample (Table 2).

The processes by which attribution to aging operate need further study. Table 3 suggests that aging attributions may occur when individuals report good health despite functional limitations. If individuals retain positive subjective assessments despite reporting physical impairments, old age may be a convenient and socially accepted explanation for the decline even though it is not specific. Other possibilities include less effective self-care, adherence to treatment, health-promoting actions,[17] or perhaps denial of the severity of problems. Whatever the processes, however, family caregivers and health care professionals should attend to comments indicating a belief that old age is the cause of health problems. Otherwise, the expected negative trajectory of aging may become a self-fulfilling prophecy with unfortunate consequences. □

TABLE 3—Bivariate Associations between Other Variables[a] and Attribution of ADL Impairment to Aging

Associated Variable	% Attributing to Aging	Odds Ratio	95% CI
Age group at baseline			
70–74	2.0	(reference category)	
75–79	3.8	1.88	.84, 4.24
80–84	5.7	2.88	1.30, 6.38
85+	14.4	8.06	3.86, 16.82
Activity compared with that of peers			
Much more	11.5	(reference category)	
More	5.4	.43	.19, .99
Same	5.3	.43	.24, .78
Less	2.7	.21	.08, .58
Much less/DK	2.8	.22	.09, .52
Reports getting enough exercise			
Yes	7.0	(reference category)	
No	4.1	.57	.36, .92
Health now compared with a year ago			
Same	6.6	(reference category)	
Better	3.8	.56	.23, 1.32
Worse/DK	3.5	.51	.29, .90
Worry due to health in past year			
None	7.0	(reference category)	
Hardly Any	8.2	1.17	.58, 2.38
Some	3.5	.48	.27, .87
A lot/DK	4.3	.59	.30, 1.18
Body mass index (quintile)			
Highest	3.2	(reference category)	
Moderately High	3.3	1.04	.43, 2.50
Middle	6.7	2.17	1.02, 4.63
Moderately Low	7.2	2.36	1.08, 5.15
Lowest	6.7	2.17	1.03, 4.55
Circulatory condition			
No	6.8	(reference category)	
Yes	4.3	.60	.38, .97

[a]Variables not associated with aging attribution included perceived control over future health, efficacy in taking care of health, frequency of trouble remembering things and experiencing confusion, number of medical visits in past year, recency of medical visit, sex, instrumental ADLs, social involvements, race, education, living alone or with others.

Acknowledgment

This research was supported in part by Agency for Health Care Policy and Research Grant R01-HS06795.

References

1. Rowe JW, Kahn RL. Human aging: usual and successful. *Science.* 1987;237:143–149.
2. Schneider EL, Reed JD Jr. Life extension. *N Engl J Med.* 1985;312:1159–1168.
3. Olshansky SJ, Carnes BA, Cassel C. In search of Methuselah: estimating the upper limits to human longevity. *Science.* 1990; 250;634–640.
4. US Department of Health and Human Services. *Healthy People 2000: National Health Promotion and Disease Prevention Objectives.* Washington, DC: US Government Printing Office; September 1990. DHHS publication PHS 91–50213.
5. Berg RL, Cassells JS, eds. *The Second Fifty Years: Promoting Health and Preventing Disability.* Washington, DC: National Academy Press; 1990.
6. Abdellah FG, Moore SR, eds. *Surgeon General's Workshop: Health Promotion and Aging, Proceedings.* Washington, DC: US Government Printing Office; 1988. US Dept. of Health and Human Services, Public Health Service.
7. Leventhal EA, Prohaska TR. Age, symptom interpretation, and health behavior. *J Am Geriatr Soc.* 1986;34:185–191.
8. Gjorup T, Henricksen C, Lund E, Stromgard E. Is growing old a disease? A study of the attitudes of elderly people to physical symptoms. *J Chron Dis.* 1987;40:1095–1098.
9. Brody EM, Kleban MH. *Mental and Physical Health Practices of Older People.* New York, NY: Springer; 1985.
10. Fitti JE, Kovar MG. The supplement on aging to the 1984 National Health Interview Survey. *Vital and Health Statistics.* Series 1, No. 21. Hyattsville, Md: National Center for Health Statistics; June 1987. US Dept of Health and Human Services publication PHS 87–1323.
11. Office of Vital and Health Statistics Systems. *The Longitudinal Study of Aging Public Use File: Version 3.* Hyattsville, Md: National Center for Health Statistics; October 1990.
12. Kovar MG. Collecting health data from and about older people: the Longitudinal Study of Aging. In: Fowler FJ Jr, ed. *Health Survey Research Methods: Conference Proceedings.* Rockville, Md: National Center for Health Services Research and Health Care Technology Assessment; September 1989. US Dept of Health and Human Services publication PHS 89–3447.
13. Katz S, Ford AB, Moskowitz RW, et al. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychological function. *JAMA.* 1963; 185:94–99.
14. Weiner JM, Hanley RJ, Clark R, Van Nostrand JF. Measuring the activities of daily living: comparison across national surveys. *J Gerontol Soc Sci.* 1990;45:S229–237.
15. Kart C. Experiencing symptoms: attribution and misattribution of illness among the aged. In: Haug M, ed. *Elderly Patients and Their Doctors.* New York, NY: Springer; 1981:70–78.
16. Idler EL, Kasl SV. Health perceptions and survival: do global evaluations of health status really predict mortality? *J Gerontol Soc Sci.* 1991;46:S55–65.
17. Prohaska TR, Keller ML, Leventhal EA, Leventhal H. Impact of symptoms and aging attribution on emotions and coping. *Health Psychol.* 1987;6:495–514.

Use of Cardiovascular Disease Medications and Mortality in People with Older Onset Diabetes

Barbara E. K. Klein, MD, MPH, Scot E. Moss, MA, and Ronald Klein, MD, MPH

ABSTRACT

Mortality follow-up for a cohort defined in 1980 with diabetes diagnosed at 30 years of age or older has been completed through 1988. History of medication use was obtained during the initial evaluation. There were 605 (44.2%) confirmed deaths; heart disease was the underlying cause in 49.9% of the deaths. Use of loop diuretics was associated with an odds ratio of death of 1.8 (95% CI = 1.4, 2.2). Although a causal relationship cannot be inferred, it is reasonable to suggest that blood chemistries be monitored regularly in persons on these drugs since electrolyte imbalance may be related to death. (*Am J Public Health.* 1992;82: 1142–1144)

Introduction

Diabetes mellitus increases the risk of death in several populations in the United States.[1–6] Much of the increase is attributed to cardiovascular disease. People with diabetes are more likely to have cardiovascular disease, including high blood pressure.[7,8] Thus, in addition to therapy for their diabetes, these people may also be taking medication to lower blood pressure and for other cardiovascular problems. Warram has reported excess mortality associated with diuretic usage in patients followed at the Joslin Clinic.[9] We explore the relationship between use of cardiovascular medications and mortality in people with diabetes diagnosed after 30 years of age in the population participating in the Wisconsin Epidemiologic Study of Diabetic Retinopathy.

Methods

The population has been described in previous reports.[10–12] Briefly, 10 135 diabetic persons were identified from the records of physicians providing primary care in an 11-county area in southern Wisconsin from July 1, 1979, through June 30, 1980. A probability sample of persons diagnosed as having diabetes after age 30 (n = 1780) was invited to the examination phase of the study from 1980 to 1982 (1370 participated). Pertinent parts of the examination included measurements of weight, height, blood pressure,[13] glycosylated hemoglobin,[14] and gross urine protein (Labstix). An interview included questions about medications for control of hyperglycemia and blood pressure and diuretic agents. We asked for the number of aspirin used in the 30 days before the baseline examination, and about smoking, angina, myocardial infarction, and stroke. Questions about medication usage were verified by a physician's report.

Participants, relatives, and contact persons have been called annually to determine vital status. Newspaper obituaries were reviewed daily. Annually, a request was made to the Wisconsin Center for Health Statistics, Section of Vital Statistics, for death certificate information for these persons. Names of persons who had been lost to follow-up were submitted for matching against death records. Wisconsin death records through 1988 have been searched. Information on persons who have moved out of Wisconsin and were suspected of being deceased and on those lost to follow-up was submitted to the National Death Index for matching against national death data.

The time interval of interest for this study ranges from the examination date through the date of death or December 31, 1988. Only deaths confirmed by a death certificate are included in the analyses. Cause-specific mortality analyses are based on the underlying cause of death or any mention on the death certificate.

People were classified as having a cardiovascular disease at examination if they reported that they had angina or had had a heart attack or stroke that was verified by a physician. People were classified as having never smoked if they had smoked fewer than 100 cigarettes in their lifetime. Proteinuria was classified as a urine protein concentration of 0.30 g/l or more as measured by Labstix.

The Cox proportional hazards model was used to evaluate the effect on mortality of a medication while controlling for other risk factors.[15] The Statistical Analysis System was used for estimating the model parameters.[16]

Results

There were 605 (44.2%) confirmed deaths (Table 1). Heart disease (codes

The authors are with the Department of Ophthalmology, University of Wisconsin, Madison.

Requests for reprints should be sent to Barbara E. K. Klein, MD, MPH, Department of Ophthalmology, University of Wisconsin, Madison, 600 Highland Avenue, Madison, WI 53792.

This paper was submitted to the Journal August 19, 1991, and accepted with revisions January 6, 1992.

TABLE 1—Frequency of Underlying Cause of Death and Mentions of Specific Causes

	ICD-9 Codes	Underlying Cause[a] n	Underlying Cause[a] %	Any Mention[a] n	Any Mention[a] %
Cancer	140–208	53	8.8	80	13.2
Diabetes	250.0–250.3, 250.5–250.9	86	14.2	361	59.7
Heart disease	391–398, 402, 404–429	302	49.9	439	72.6
Ischemic heart	410–414	237	39.2	297	49.1
Other heart	391–398, 402, 404, 415–429	65	10.7	330	54.5
Stroke	430–438	52	8.6	108	17.9
Renal	250.4, 580–589	8	1.3	66	10.9
Other	...	104	17.2	...	...
All causes	...	605	100.0	...	...

Note. ICD-9 = *International Classification of Diseases*, 9th version.
[a]Sums of sample size and percentage exceed all causes because multiple causes may be mentioned on the death certificate.

TABLE 3—Odds of Death Associated with Use of Specific Medications

Causes and Medication[a]	Underlying Cause P	Underlying Cause Odds Ratio[b]	Underlying Cause (95% CI)	Any Mention P	Any Mention Odds Ratio[b]	Any Mention (95% CI)
All causes						
Loop diuretics	<.0001	1.8	(1.4, 2.2)	...	...	...
Cancer						
Beta blockers	.01	2.9	(1.2, 6.8)	...	...	...
Tolbutamide	.03	0.2	(0.0, 1.1)	.01	0.2	(0.1, 0.8)
Acetohexamide	...	...	...	.05	2.2	(1.0, 4.7)
Diabetes						
Loop diuretics	<.01	2.3	(1.2, 4.3)	<.0001	1.8	(1.4, 2.5)
Heart disease						
Beta blockers	...	...	...	.04	1.4	(1.0, 2.0)
Loop diuretics	<.0001	2.0	(1.5, 2.7)	<.0001	2.0	(1.5, 2.6)
Ischemic heart disease						
Loop diuretics	<.0005	2.0	(1.4, 2.8)	<.0001	1.9	(1.4, 2.6)
Other heart disease						
Loop diuretics	<.01	2.3	(1.2, 4.4)	<.0001	2.6	(2.0, 3.5)
Thiazide diuretics	.05	1.7	(1.0, 3.0)	...	...	...
Stroke						
Thiazide diuretics	<.001	2.8	(1.5, 5.3)	<.01	1.8	(1.2, 2.7)
Renal						
Vasodilators	...	...	...	<.005	3.7	(1.6, 8.5)
Loop diuretics	...	...	...	<.0001	3.7	(1.9, 7.0)
All other causes						
K-sparing diuretics	.03	1.9	(1.1, 3.4)	...	...	...

[a]Medications considered: aspirin, beta blockers, centrally and peripherally acting anti-adrenergic agents, vasodilators, K-sparing diuretics, loop diuretics, thiazide diuretics, acetohexamide, chlorpropamide, tolbutamide, and tolazamide.
[b]Controlling for age, sex, systolic blood pressure, presence of proteinuria, history of cardiovascular disease, history of cancer (for cancer as a cause of death only), smoking history, body mass, duration of diabetes, glycosylated hemoglobin, and insulin use.

391–398, 402, and 404–429 of the *International Classification of Diseases*, 9th version) was listed as the underlying cause of 49.9% of deaths. Table 2 lists the frequencies of antihypertensive and diuretic medications, aspirin, and tolbutamide use. Many subjects took more than one of these medications.

The odds of death for use of these medications were computed. Other relevant characteristics included in the analyses were age, duration of diabetes, use of insulin (no or yes), systolic and diastolic blood pressures, glycosylated hemoglobin (as a percentage of hemoglobin), sex, body mass index (weight/height^2), smoking status (current, past, or never smoked), and presence of proteinuria. The latter was included in an attempt to control for the presence of renal disease. The odds of death as derived from proportional hazards analyses appear in Table 3. The use of loop diuretics was associated with all causes of mortality and with mortality attributed to diabetes, heart disease, and ischemic heart disease. Use of thiazide diuretics was associated with death due to heart disease, other than ischemic, and with stroke. Potassium sparing diuretics were associated with all other causes of mortality. These analyses were repeated after excluding those who had proteinuria. There was no change in the pattern of significant associations between use of loop diuretics and death. Also, major drug combinations were examined. These included use of a beta blocker or potassium sparing diuretic with a thiazide diuretic and use of a beta blocker or potassium sparing diuretic with a loop diuretic. In all cases, except for the association of loop diuretics with death due to other heart disease, the effect of loop and thiazide diuretics remained significant after controlling for use of beta blockers or potassium sparing diuretics and interactions with these agents.

TABLE 2—Frequencies of Reported Use of Selected Drugs in the Cohort

Drug	n	%
Alpha blockers	23	1.7
Beta blockers	120	8.8
Loop diuretics	177	12.9
Thiazide diuretics	552	40.3
Potassium sparing diuretics	176	12.8
Vasodilators	58	4.2
Aspirin	338	25.0
Tolbutamide	117	8.5

Discussion

Patients with diabetes mellitus are at increased risk of death due to cardiovascular diseases.[17] One might anticipate that the presence of these diseases would necessitate the use of any of the medications of interest in this paper. Therefore, in order to attempt to control for this "confounding by indication," we included a history of cardiovascular disease in our analyses. For each cause of death cate-

gory, history of cardiovascular disease was a significant variable. However, use of at least one diuretic agent added significant information for most of the models. Loop agents were the agents most frequently included, but thiazide diuretics were also significantly associated with an increased risk of death. The differences between the analyses indicating one class of diuretics as important compared with another may be a reflection of the frequency of their use rather than the particular drug effect. In clinical practice, nephrologists may avoid using thiazides because of their effects on glycemia and uricemia. There are some studies that have indicated reduced mortality and morbidity in hypertensive subjects who receive medical treatment,[18] but an association between diuretic use and arrhythmias has been reported.[19–21] Whether the use of these agents with diabetic subjects truly predisposes to death cannot be determined from these analyses. However, it is probably wise for clinicians caring for such patients to be especially attentive to blood chemistries. Also, it has been suggested that the use of diuretics in patients with diabetic nephropathy may accelerate that condition.[22]

In these analyses we evaluated a possible relationship of mortality to use of tolbutamide. This drug had been implicated as a cause of increased mortality in the University Group Diabetes Program.[23] Its use was relatively infrequent in our study population (Table 2); it did not contribute significantly to describing mortality. □

Acknowledgments

This research was supported by National Institutes of Health–National Eye Institute Grant EY 03083 to Ronald Klein.

We are grateful to Fred Krantz, Wisconsin Center for Health Statistics, for providing death certificate information on Wisconsin residents; to Betty Sullivan and Peggy Peterson, Wisconsin Center for Health Statistics, for coding certificates of out-of-state residents; to Stacy Meuer and Ane Weber for data processing; and to Julie K. Olson and Kristine A. Tway for assistance with the manuscript.

References

1. Moss SE, Klein R, Klein BEK. Cause-specific mortality in a population-based study of diabetes. *Am J Public Health.* 1991;81:1158–1162.
2. Kessler II. Mortality experience of diabetic patients. A twenty-six year follow-up study. *Am J Med.* 1976;51:715–724.
3. Palumbo PJ, Elveback LR, Chu C-P, Connolly DC, Kurland LT. Diabetes mellitus: incidence, prevalence, survivorship, and causes of death in Rochester, Minnesota, 1945–1970. *Diabetes.* 1976;25:566–573.
4. Herman JB, Medalie JH, Goldbourt U. Differences in cardiovascular morbidity and mortality between previously known and newly diagnosed adult diabetes. *Diabetologia.* 1977;13:229–234.
5. Krolewski AS, Czyzy KA, Janeczko D, Kopczynski J. Mortality from cardiovascular disease among diabetics. *Diabetologia.* 1977;13:345–350.
6. Kannel WB, McGee DL. Diabetes and glucose tolerance as risk factors for cardiovascular disease: The Framingham Study. *Diabetes Care.* 1979;2:120–126.
7. Klein R, Klein BEK, Moss SE, DeMets DL. Blood pressure and hypertension in diabetes. *Am J Epidemiol.* 1985;122:75–89.
8. Dupree EA, Meyer MB. Role of risk factors in complications of diabetes mellitus. *Am J Epidemiol.* 1980;112:100–112.
9. Warram JH, Laffel LMD, Valsania P, Christleib AR, Krolewski AS. Excess mortality associated with diuretic therapy in diabetes mellitus. *Arch Intern Med.* 1991; 151:1350–1356.
10. Klein R, Klein BEK, Moss SE, DeMets DL, Kaufman I, Voss PS. Prevalence of diabetes mellitus in southern Wisconsin. *Am J Epidemiol.* 1984;119:54–61.
11. Klein R, Klein BEK, Moss SE, Davis MD, DeMets DL. The Wisconsin Epidemiologic Study of Diabetic Retinopathy II. Prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 years. *Arch Ophthalmol.* 1984;102:520–526.
12. Klein R, Klein BEK, Moss SE, Davis MD, DeMets DL. The Wisconsin Epidemiologic Study of Diabetic Retinopathy III. Prevalence and risk of diabetic retinopathy when age at diagnosis is 30 or more years. *Arch Ophthalmol.* 1984;102:527–532.
13. Hypertension Detection and Follow-up Program Cooperative Group. The hypertension detection and follow-up program. *Prev Med.* 1976;5:207–215.
14. Isolab. *Quick Step, Fast Hemoglobin Test System.* Akron, Ohio: Isolab; 1981:1–8.
15. Cox DR. Regression models and life tables (with discussion). *J Royal Stat Soc B.* 1972; 34:187–220.
16. SAS Institute Inc. *SUGI Supplemental Library User's Guide Version 5 Edition.* Cary, NC: SAS Institute Inc; 1986:437–466.
17. Kleinman JC, Donahue RP, Harris MI, Finucane FF, Madans JH, Brock DB. Mortality among diabetics in a national sample. *Am J Epidemiol.* 1988;128:389–401.
18. Hartford M, Wendelhad I, Berglund G, Wallentin I, Ljungman S, Wikstrand J. Cardiovascular and renal effects of long-term antihypertensive treatment. *JAMA.* 1988;259:2553–2557.
19. Hollan OB, Nixon JV, Kuhnert L. Diuretic-induced ventricular ectopic activity. *Am J Med.* 1981;70:762–768.
20. Kuller LH, Hulley SB, Cohen JD, Neaton J. Unexpected effects of treating hypertension in men with electrocardiographic abnormalities: a critical analysis. *Circulation.* 1986;73:114–123.
21. Caralis PV, Perez-Stable E. Electrolyte abnormalities and ventricular arrhythmia. *Drugs.* 1986;31(suppl 4):85–100.
22. Walker WG, Hermann J, Yin D, Murphy RP, Patz A. Diuretics accelerate diabetic nephropathy in hypertensive insulin-dependent and non-insulin-dependent subjects. *Trans Assoc Am Physicians.* 1987; 100:305–315.
23. University Group Diabetes Program. Mortality results. *Diabetes.* 1970;19(suppl 2):789–830.

The Accuracy of the National Death Index When Personal Identifiers Other than Social Security Number Are Used

Brent C. Williams, MD, MPH, Lucy B. Demitrack, BA, and Brant E. Fries, PhD

ABSTRACT

This study analyzed the accuracy of the National Death Index when personal identifiers were used that included or excluded Social Security number. Computerized records of the Department of Veterans Affairs were used for comparison. Different combinations of identifiers other than Social Security number correctly identified from 83 to 92 percent of dead and 92 to 99 percent of living persons. These results should prove useful in ascertaining the mortality status of patient populations without information on Social Security numbers. (*Am J Public Health.* 1992;82:1145–1147)

Introduction

Accurate determination of mortality among subjects is critical for many types of research. To meet this need, the National Death Index has been established as a national computerized database of death records compiled by the National Center for Health Statistics.

Several studies have examined the accuracy of the National Death Index,[1–10] but all were completed during its first few years of operation, and few have focused on the comparative accuracy of different combinations of information that include or exclude Social Security number in correctly establishing death. Yet, for some populations information on Social Security number is not readily available. The usefulness of the National Death Index in determining the vital status of such groups depends on the accuracy of the matching process using information other than Social Security number.

We had the opportunity to compare the accuracy of the National Death Index using information that includes or excludes Social Security number during a study of health services use among users of facilities of the Department of Veterans Affairs.

Methods

The National Death Index

Details of the National Death Index application and matching process have been described elsewhere.[1,3,5] Briefly, the index contains information from all death certificates on file in state vital statistics offices since 1979, including decedents' names, Social Security numbers, dates of birth, and several demographic characteristics. A National Death Index user may specify which year(s) of death should be searched. Records are available for searches approximately 12 to 15 months after the end of a given calendar year. After approval and upon receipt of identifying information from the investigator, all cases from the National Death Index files that meet one or more of 12 selection criteria are appended and returned to the investigator for evaluation. The selection criteria have been described elsewhere[3] (and are available from National Death Index, Division of Vital Statistics, National Center for Health Statistics, 6525 Belcrest Road, Room 840, Hyattsville, MD 20782).

Data Sources and Variables

Data for our submission to the National Death Index were compiled from computerized hospital and nursing home records of all residents of Department of Veterans Affairs nursing homes between October 1986 and April 1987. Among other data requested, information for two groups of persons was submitted: (1) a random sample of persons recorded as dying in Department of Veterans Affairs hospitals or nursing homes through October 1, 1988 ("known dead," n = 500), and (2) all persons discharged alive or who received a clinical assessment in a Department of Veterans Affairs nursing home between September 15, 1988, and October 1, 1988 ("known alive," n = 3487). These data were matched with records from the National Death Index for the years 1986 through 1988, excluding index records with dates of death after October 15, 1988.

Analytic Methods

Department of Veterans Affairs records were assumed to reflect accurately patient mortality status. The analysis was

Brent C. Williams and Brant E. Fries are with the Ann Arbor Veterans Administration Medical Center Geriatrics Research Education and Clinical Center; Brent C. Williams is also with the Department of Internal Medicine, and Brant E. Fries with the Institute of Gerontology and School of Public Health, University of Michigan, Ann Arbor. Lucy B. Demitrack is with the Institute of Gerontology and School of Public Health, University of Michigan.

Requests for reprints should be sent to Brent C. Williams, MD, MPH, University of Michigan Medical Center, Department of Internal Medicine, Division of General Medicine, 3116 Taubman Center, Ann Arbor, MI 48109-0376.

This paper was submitted to the Journal October 27, 1991, and accepted with revisions April 21, 1992.

TABLE 1—Accuracy of Patient Identifiers in Ascertaining Death from the National Death Index

Identifiers	Identifiers Alone, %		Identifiers and Social Security Number, %	
	Sensitivity[a] (n = 500)	Specificity[b] (n = 3487)	Sensitivity[a] (n = 500)	Specificity[b] (n = 3487)
FI, LN, DB, MB, YB	83	99	83	99
FI, LN, DB, MB	85	97	85	99
FI, LN, MB, YB	88	93	88	99
FI, LN, YB	90	93	89	99
FI, LN, YB or LN, DB, MB, YB	92	92	92	99
Social Security number	...	...	97	99

Note. FI = first initial; LN = last name; DB = day of birth; MB = month of birth; YB = year (±1 year) of birth.
[a]Percentage of persons "known dead" with one or more National Death Index records that match on specified criteria.
[b]Percentage of persons "known alive" with no National Death Index records that match on specified criteria.

designed to determine the "sensitivity" and "specificity" of the National Death Index for ascertaining patient mortality. Sensitivity refers to the percentage of records of persons who were known dead that had at least one matching National Death Index record for a given set of identifiers. Specificity refers to the percentage of records of persons known alive that had no matching National Death Index records for a given set of identifiers. Sensitivity and specificity are collectively referred to as "accuracy."

A match on Social Security number was defined to occur if eight or nine digits were identical in the Department of Veterans Affairs and National Death Index records. A match on year of birth occurred if the two records differed by at most 1 year, an assumption that has been shown to improve the accuracy of matching.[5]

Results

Table 1 shows the accuracy of several combinations of identifiers in ascertaining the mortality status of members of our population, including or excluding Social Security number. Based on name and birth date alone, from 83% to 92% of known decedents and 92% to 99% of persons known to be alive were identified. The matching combination with the highest sensitivity (92%) was either (1) first initial, last name, and year of birth or (2) last name and complete (day, month, and year) date of birth. For investigators without Social Security number information, this combination would identify the greatest proportion of deaths with an acceptably high specificity (92%).

Among the known dead, the vast majority of "true" matches based on name and date of birth also matched on Social Security number. Thus, including information on Social Security number did not change the sensitivities substantially. However, of the observations from the known alive group that matched on name and birth-date identifiers, a substantial portion did not match on Social Security number. This was due to a relatively large number (n = 131) of (probably false) matches only on first initial, last name, and month and year of birth, and not on Social Security number. Including Social Security number as a matching criterion therefore significantly decreased the number of false-positive matches.

Social Security number alone had the best accuracy of any combination of identifiers examined, with a sensitivity of 97% and a specificity of 99%. The improved sensitivity of Social Security number alone over any other combination of name and birth-date identifiers was largely due to 13 cases among the known dead that (probably truly) matched on Social Security number and exact birth date, but not on last name. Matching on Social Security number alone thus yielded the highest number of true matches.

Adding information on race and/or marital status to the optimal combination of variables based on name and birth date did not significantly improve the overall results.

Discussion

These results demonstrate that among patient identifiers other than Social Security number, either first initial, last name, and year of birth or last name and full date of birth provides optimal ascertainment of death using the National Death Index. Matching on Social Security number alone, however, provided the best overall results. Therefore, investigators who wish to use the National Death Index to determine the vital status of a study population should include information on Social Security number whenever possible. If this is not possible, however, the National Death Index remains a highly accurate and useful resource and will correctly identify the mortality status of at least 92% of individuals using the criteria applied in this study.

Our estimates of the sensitivity and specificity of the National Death Index using identifiers other than Social Security number are probably conservative, for two reasons. First, inevitable misclassification of patients' mortality status in the Department of Veterans Affairs data would result in falsely low estimates. Second, the Department of Veterans Affairs data include only first initial rather than first name; complete information on first name would probably improve the specificity of the results without substantially decreasing the sensitivity.

This study had two important limitations: our data included few women and did not allow stratification of the results by race. Other studies[2,8] have suggested that the National Death Index may be less accurate for women and racial minorities than for other groups. The generalizability of these results to women and minorities should therefore be evaluated in future studies. □

References

1. Wentworth DN, Neaton JD, Rasmussen WL. An evaluation of the Social Security Administration Master Beneficiary Record File and the National Death Index in the ascertainment of vital status. *Am J Public Health.* 1983;73:1270–1274.
2. Curb JD, Ford CE, Pressel S, Palmer M, Babcock C, Hawkins CM. Ascertainment of vital status through the National Death Index and the Social Security Administration. *Am J Epidemiol.* 1985;121:754–766.
3. Edlavitch SA, Baxter J. Comparability of mortality follow-up before and after the National Death Index. *Am J Epidemiol.* 1988;127:1164–1178.
4. Boyle CA, Decouflé P. National sources of vital status information: extent of coverage and possible selectivity in reporting. *Am J Epidemiol.* 1990;131:160–168.
5. Patterson JE, Bilgrad R. The National Death Index experience: 1981–1985. In: Kilss B, Albey W, eds. *Proceedings of the Workshop on Exact Matching Methodologies, Arlington, VA, May 9–10, 1985.* Washington, DC: US Government Printing Office; 1985:245–254. US Department of the Treasury, Internal Revenue Service, publication no. 1299 (2-86).

6. Rogot E, Feinleib M, Ockay KA, Schwartz SH, Bilgrad R, Patterson JE. On the feasibility of linking census samples to the National Death Index for epidemiologic studies: a progress report. *Am J Public Health.* 1983;73:1265–1269.
7. Rogot E, Sorlie P, Johnson NJ. Probabilistic methods in matching census samples to the National Death Index. *J Chronic Dis.* 1986;39:719–734.
8. Acquavella JF, Donaleski D, Hanis NM. An analysis of mortality follow-up through the National Death Index for a cohort of refinery and petrochemical workers. *Am J Ind Med.* 1986;9:181–187.
9. Davis KB, Fisher L, Gillespie MJ, Pettinger M. A test of the National Death Index using the Coronary Artery Surgery Study (CASS). *Controlled Clin Trials.* 1985;6:179–191.
10. Stampfer MJ, Willett WC, Speizer FE, et al. Test of the National Death Index. *Am J Epidemiol.* 1984;119:837–839.

Race and Sex Differences in Mortality following Fracture of the Hip

Steven J. Jacobsen, MD, PhD, Jack Goldberg, PhD, Toni P. Miles, MD, PhD, Jacob A. Brody, MD, William Stiers, PhD, and Alfred A. Rimm, PhD

ABSTRACT

This paper describes the all-cause mortality experience, following a fracture of the hip, of 712 027 persons covered by the Medicare program from 1984 through 1987. White women experienced the lowest mortality rate (17.2 per 1000 person-months), followed by Black women (22.9 per 1000 person-months), Black men (33.5 per 1000 person-months), and White men (33.7 per 1000 person-months). The observed race-sex differences in survival were found at all ages and regardless of the number of comorbid conditions listed with the discharge diagnosis. While these data demonstrate marked race-sex differences in survival following hip fracture, the cause of these differences is not immediately apparent and demands further investigation. (*Am J Public Health.* 1992;82:1147–1150)

Introduction

Fracture of the hip represents a tremendous burden to the population aged 65 years and older.[1] White women experience the highest incidence of fracture, followed by White men, Black women, and Black men.[2–4] Following fracture, persons are at a much greater risk of death compared to the general population.[5–17] However, most studies of survival have been conducted in Scandinavia[6,11–13] and other predominately White populations,[14–16] and it is not known how well they generalize to the entire United States population or to minorities. The few studies that have included Blacks have been severely limited by a small sample size. The purpose of this study is to provide stable estimates of the survival following hip fracture of more than 700 000 Whites and Blacks aged 65 years and older in the United States.

Methods

Data for this study were obtained from the Health Care Financing Administration (HCFA) and have been described in detail elsewhere.[4,18,19] Briefly, HCFA maintains a computerized database of all hospital discharges covered by the Medicare program in the Medicare Provider Analysis and Review (MedPAR) file.[20] These data, keyed by Social Security Number, include the patient's age, race, and sex; up to five discharge diagnoses; zip code of residence; and dates of admission and discharge. To ascertain a date of death, records are cross-tabulated against death records from the Social Security Administration on a quarterly basis for two years following the date of admission; this follow-up information is incorporated in the summary file.

In the 1984 through 1987 MedPAR files, potential cases of hip fracture were identified as any individual with an ICD9-CM code[21] of 820.0 through 820.9 in any of the five fields provided for diagnoses (n = 810 949). Cases were excluded if age was less than 65 years (n = 19 899), if the fracture was the second or subsequent hip fracture during the four years (n = 54 055), if race was coded as other or unknown (n = 24 269), or if the fracture may have been secondary to neoplastic disease or the primary discharge diagnosis was for late effects of hip fracture (n = 4464). After exclusions, a total of 712 027 cases of hip fracture remained. (Some cases were excluded for more than one reason.)

The survival experience of persons following fracture of the hip was quantified

Steven J. Jacobsen is with the Department of Health Sciences Research, Section of Clinical Epidemiology, Mayo Clinic, Rochester, Minn. Jack Goldberg, Toni P. Miles, and Jacob A. Brody are with the Epidemiology Program, School of Public Health, University of Illinois at Chicago; Jack Goldberg is also with the VA Cooperative Studies Program Coordinating Center, Hines Hospital, Hines, Ill. William Stiers is with the Division of Rehabilitation Psychology and Neuropsychology, Department of Physical Medicine and Rehabilitation, University of Michigan Medical Center, Ann Arbor, Mich. Alfred A. Rimm is with the Division of Biostatistics/Clinical Epidemiology, Medical College of Wisconsin, Milwaukee, Wis.

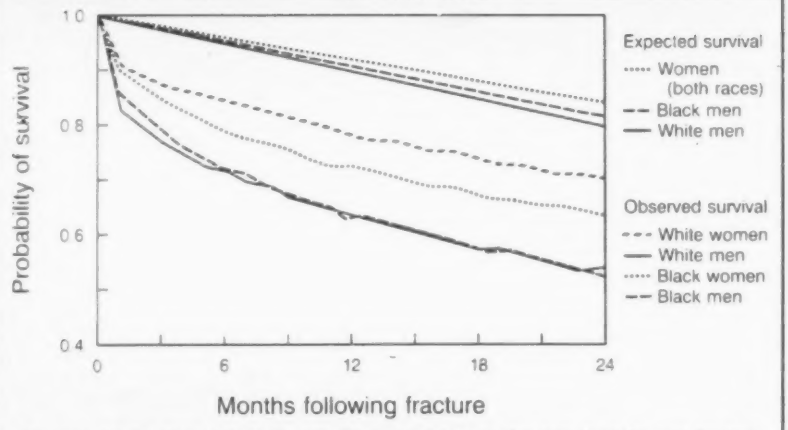
Requests for reprints should be sent to Steven J. Jacobsen, MD, PhD, Department of Health Sciences Research, 200 First St SW, Mayo Clinic and Foundation, Rochester, MN 55905.

This paper was submitted to the Journal December 17, 1991, and accepted with revisions May 27, 1992.

TABLE 1—Mortality Experience of Hip Fracture Patients Aged 65 Years and Older, 1984 to 1987

	n	Deaths	Follow-up (Person-Months)	Mortality Rate (per 1000)
White Men				
65–74	36 473	10 223	540 291	18.9
75–84	62 513	25 664	792 639	32.4
85–94	40 975	21 926	432 168	50.7
95+	4088	2736	32 371	84.5
Total	144 049	60 549	1 797 469	33.7
White Women				
65–74	103 015	15 957	1 703 750	9.4
75–84	233 047	51 890	3 623 108	14.3
85–94	189 459	63 790	2 619 480	24.4
95+	18 247	8863	201 990	43.9
Total	543 768	140 500	8 148 328	17.2
Black Men				
65–74	2295	670	33 956	19.7
75–84	2902	1228	35 852	34.3
85–94	1659	932	16 584	56.2
95+	208	126	1735	72.6
Total	7084	2956	88 127	33.5
Black Women				
65–74	3425	722	53 185	13.6
75–84	6819	1957	97 106	20.2
85–94	5968	2284	76 070	30.0
95+	934	449	9842	45.6
Total	17 146	5412	236 203	22.9

Source. Health Care Financing Administration Medicare Provider Analysis and Review Record: 1984–1987[20].



Sources. Health Care Financing Administration Medicare Provider Analysis and Review Record: 1984–1987[20]; 1980 Decennial Life Table Estimates.[25]

FIGURE 1—Observed and expected race- and sex-specific survival following fracture of the hip, all ages combined.

using mortality rates and survival function estimates. Follow-up (measured in person-months) was started on the date of admission and continued until death or censoring (January 1 of the second year following admission or 24 months, whichever was shorter). Mortality rates were calculated by dividing the number of deaths within a stratum by the stratum-specific person-time denominator[22]; survival functions were estimated using the Kaplan-Meier method.[23] Expected mortality rates within sex- and race-specific stratum were calculated[24] using the 1-year age-specific 1980 Decennial Life Table Estimates.[25]

Results

During the study period there were 687 817 discharges for hip fracture among Whites and 24 210 among Blacks (Table 1), accounting for a total of 10 270 127 person-months of follow-up. Men experienced the highest mortality, with nearly identical rates among Blacks and Whites. Rates among women were markedly lower than men, with White women experiencing lower mortality rates than Black women. The survival curves for Black women and White women diverged from the first through ninth month following fracture, thereafter remaining parallel (Figure 1). Regardless of race, men fared much worse than women.

Among all four race- and sex-specific strata, survival during the first two months following fracture was much worse than we expected, whereas from six months and onward the survival curves roughly paralleled expectations. Although mortality increased in successively older age groups among all races, the relative position of the race- and sex-specific survival curves remained constant across all age strata (Figure 2). This was also true when survival curves were stratified by the number of comorbid conditions[26] (Figure 3; see Appendix for list of comorbid conditions).

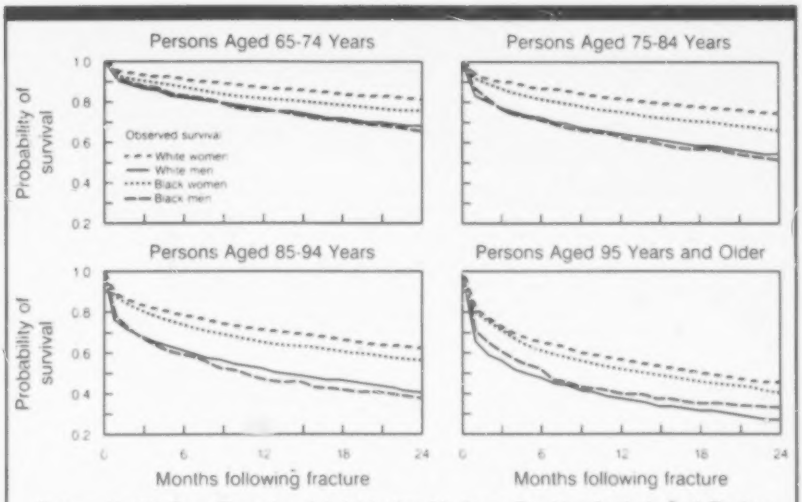
Discussion

In this study we report the survival of 712 027 Whites and Blacks aged 65 years

and older who suffered a hip fracture. Mortality was the greatest during the first 2 months following fracture but paralleled that of the general population after 6 to 9 months. In general, women survived longer than men and Black women fared relatively worse than their White counterparts. Two studies[15,17] reporting survival among Black women found them to fare worse than White women, whereas one reported better survival.[16] However, none of these studies could examine racial differences among men.

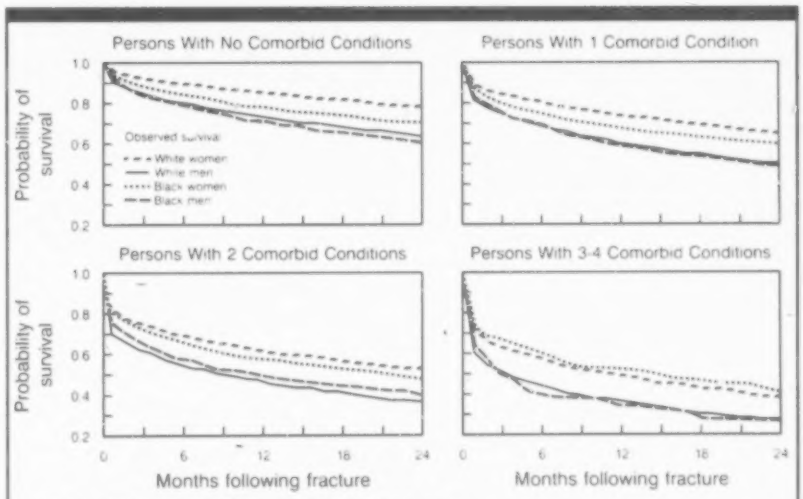
Although the observed differences are striking, potential biases must be considered when interpreting these data. The rates for Whites are consistent with previously reported results both within the United States[5,8-10,14-17] and in Europe.[6,11-13] Because race is determined prospectively, systematic miscoding of race conditional on longevity seems highly unlikely. Further, under-reporting of deaths among Blacks would tend to close the racial gap in survival. A lag in reporting time is an unlikely explanation for the differences between Blacks and Whites because the relative position of the race- and sex-specific survival curves is similar for the first six months of 1984 and last six months of 1986 (data not shown). The potential underascertainment of hip fractures[27] by relying solely on the Medicare Part A data would not explain the differences unless there was a differential under-reporting on the basis of both race and survival, which seems unlikely.

Because bias does not seem to be responsible for the poorer relative survival among Black women, etiologic causes must be sought. In contrast to mortality, the incidence of fracture among Blacks is lower than Whites[2-4]; this has been attributed to greater bone mass[28-31] and an increased prevalence of obesity.[32] Consequently, a Black person with a hip fracture may be sicker and lose proportionately more bone mass than a White person with a hip fracture. However, stratification by number of comorbid conditions suggests that the observed differences may not be due to differences in burden of morbidity. (This conclusion is limited by our measure of comorbidity, which included only diagnoses from the index hospitalization, weighted all conditions equally, and did not account for pre-fracture functional status.) Alternatively, race and sex differences in survival could be due to differences in medical care utilization and delivery, either during hospitalization or subsequent to discharge.[17] These possibilities need to be investigated more thoroughly. □



Source. Health Care Financing Administration Medicare Provider Analysis and Review Record: 1984–1987.[20]

FIGURE 2—Age-, race-, and sex-specific survival following fracture of the hip.



*For a list of the comorbid conditions, see Appendix.

Source. Health Care Financing Administration Medicare Provider Analysis and Review Record: 1984–1987.[20]

FIGURE 3—Race- and sex-specific survival following fracture of the hip, stratified by number of comorbid conditions.*

Acknowledgments

This research was supported by NIH grant RO-1-AG-08709 and a grant from the National Osteoporosis Foundation. A portion of this work was presented at the 23rd Annual Meeting of the Society for Epidemiologic Research, Snowbird, Utah, in June 1990.

The authors would like to thank Henry Krakauer, MD, PhD, for his help in assembling the data used for this study. Also, the editorial assistance of Elizabeth F. Dombrowski, MA, and the secretarial assistance of Mrs. Susie Goodman and Mrs. Sondra Buehler were greatly appreciated.

References

1. Cummings SR, Kelsey JL, Nevitt MC, O'Dowd KJ. Epidemiology of osteoporosis and osteoporotic fractures. *Epidemiol Rev.* 1985;7:178–208.
2. Farmer ME, White LR, Brody JA, Bailey KR. Race and sex differences in hip fracture incidence. *Am J Public Health.* 1984; 74:1374–1380.
3. Silverman SL, Madison RE. Decreased incidence of hip fracture in Hispanics, Asians, and Blacks: California hospital discharge data. *Am J Public Health.* 1988;78: 1482–1483.

4. Jacobsen SJ, Goldberg J, Miles TP, Brody JA, Stiers W, Rimm AA. Hip fractures incidence among the old and very old: a population-based study of 745,435 cases. *Am J Public Health.* 1990;80:871–873.
5. Fitts WT Jr, Lehr HB, Schor S, Roberts B. Life expectancy after fracture of the hip. *Surg Gynecol Obstet.* 1959;108:7–12.
6. Alffram PA. An epidemiologic study of cervical and trochanteric fractures of the femur in an urban population. *Acta Orthop Scand.* 1964;(suppl 65):1–109.
7. Gordon PC. The probability of death following a fracture of the hip. *Can Med Assoc J.* 1971;105:47–51.
8. White BL, Fisher WD, Laurin CA. Rate of mortality for elderly patients after fracture of the hip in the 1980's. *J Bone Joint Surg [Am].* 1987;69-A.1335–1340.
9. Miller CW. Survival and ambulation following hip fracture. *J Bone Joint Surg [Am].* 1978;60-A:930–933.
10. Katz S, Heiple KG, Downs TD, Ford AB, Scott CP. Long term course of 147 patients with fracture of the hip. *Surg Gynecol Obstet.* 1967;124:1219–1230.
11. Elmerson S, Zetterberg C, Andersson GBJ. Ten-year survival after fractures of the proximal end of the femur. *Gerontology.* 1988;34:186–191.
12. Jensen JS, Tondevold E. Mortality after hip fractures. *Acta Orthop Scand.* 1979;50: 161–167.
13. Dahl E. Mortality and life expectancy after hip fractures. *Acta Orthop Scand.* 1980;51: 163–170.
14. Gallagher JC, Melton LJ III, Riggs BL, Bergstralh E. Epidemiology of fractures of the proximal femur in Rochester, Minnesota. *Clin Orthop.* 1980;150:163–171.
15. Magaziner J, Simonsick EM, Kashner TM, Hebel JR, Kenzora JE. Survival experience of aged hip fracture patients. *Am J Public Health.* 1989;79:274–278.
16. Fisher ES, Baron JA, Malenka DJ, et al. Hip fracture incidence and mortality in New England. *Epidemiology.* 1991;2:116–122.
17. Furstenberg AL, Mezey MD. Differences in outcome between Black and White elderly hip fracture patients. *J Chronic Dis.* 1987;40:931–938.
18. Jacobsen SJ, Goldberg J, Miles TP, Brody JA, Stiers W, Rimm AA. Regional variation in the incidence of hip fracture: US White women aged 65 years and older. *JAMA.* 1990;264:500–502.
19. Jacobsen SJ, Goldberg J, Miles TP, Brody JA, Stiers W, Rimm AA. Seasonal variation in the incidence of hip fracture among White persons aged 65 years and older in the United States, 1984–1987. *Am J Epidemiol.* 1991;133:996–1004.
20. *Medicare Statistical Files Manual.* Baltimore, Md: Health Care Financing Administration, Bureau of Data Management and Strategy; 1990.
21. Breslow L, ed. *The International Classification of Diseases, 9th revision: Clinical Modification.* Hyattsville, Md: Health Care Financing Administration; 1980;1-3. DHHS publication PHS 80-1260.
22. Rothman KJ. *Modern Epidemiology.* Boston: Little, Brown, and Co; 1986.
23. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 1958;53:457–481.
24. Hartz AJ, Giefer EE, Hoffmann RG. A comparison of two methods for calculating expected mortality. *Stat Med.* 1983;2(3): 381–386.
25. *US Decennial Life Tables for 1979–81.* Hyattsville, Md: National Center for Health Statistics; August 1985;1(1). DHHS publication PHS 85-1150-1.
26. *Medicare hospital mortality information 1986.* Washington, DC: Health Care Financing Administration; 1987.
27. Fisher ES, Baron JA, Malenka DJ, Barrett J, Bubolz TA. Overcoming potential pitfalls in the use of Medicare data for epidemiologic research. *Am J Public Health.* 1990;80:1487–1490.
28. Trotter M, Broman GE, Peterson RR. Densities of White and Negro skeletons. *J Bone Joint Surg [Am].* 1960;42-A:50–58.
29. Moldawer M, Zimmerman SJ, Collins LC. Incidence of osteoporosis in elderly Whites and elderly Negroes. *JAMA.* 1965;194: 859–862.
30. Baker PT, Angel JL. Old age changes in bone density: sex and race factors in the United States. *Hum Biol.* 1965;37:104–121.
31. Mangaroo J, Glasser JH, Roht LH, Kapadia AS. Prevalence of bone demineralization in the United States. *Bone.* 1985; 6:135–139.
32. Farmer ME, Harris T, Madans JH, Wallace RB, Cornoni-Huntley J, White LR. Anthropometric indicators and hip fracture: the NHANES I epidemiologic follow-up study. *J Am Geriatr Soc.* 1989;37:9–16.

APPENDIX—Comorbid Conditions Listed with Discharge Diagnosis

Neoplastic (ICD9-CM codes 141.0 through 160.9, 162.0 through 172.9, or 174.0 through 209.9), pulmonary (480.0 through 487.9, 491.0 through 492.9, or 496.0 through 496.9), congestive heart failure (398.91, 402.01, 402.11, 402.91, 428.0 through 428.9, or 785.0), acute myocardial infarction (410.0 through 410.9), hepatic disease (571.0 through 572.9), renal disease (582.0 through 583.9, 585.0 through 587.9), psychiatric conditions/dementia (290.0 through 290.9, 294.0 through 299.9), hypertension (402.0 through 405.9), cerebrovascular disease (430.0 through 438.9), diabetes (250.01, 250.1 through 250.9), and other cardiovascular disease (412.0 through 414.9, 426.0 through 429.2).

Seroprevalence of HIV-1 and Hepatitis B and C in Prostitutes in Albuquerque, New Mexico

Stephen R. Tabet, MD, Darwin L. Palmer, MD, William H. Wiese, MD, MPH, Ronald E. Voorhees, MD, MPH, and Dorothy R. Pathak, PhD

ABSTRACT

A survey of persons soliciting sex in an area known to be frequented by prostitutes in Albuquerque, NM, included 43 females and 66 males. Seroprevalence rates found in this population-based study were as follows: human immunodeficiency virus type 1 (HIV-1), 3%; hepatitis B, 39%; hepatitis C, 45%. Increased age, intravenous drug use, and condom use were independent risk factors for hepatitis B. Female gender and intravenous drug use were independent risk factors for hepatitis C. Neither sharing injection equipment nor engaging in receptive anal intercourse was independently associated with hepatitis B or C. (*Am J Public Health.* 1992;82: 1151–1154)

Introduction

Prevalence rates of human immunodeficiency virus type 1 (HIV-1) infection in female prostitutes in the United States have been reported to be as high as 58 percent.[1,2] These prostitutes' risk of infection is probably associated more with intravenous (IV) drug use than with sexual activity.[3,4] Male and female prostitutes have high rates of hepatitis B (HBV) infection.[5] Hepatitis C (HCV) rates in prostitutes are not well known, but a high prevalence in IV drug users has been documented.[6,7] Although transmission patterns for human lymphotropic virus types I and II (HTLV I/II) are unclear, the prevalence in 1305 female prostitutes in the United States was almost 7%.[8]

Most published surveys of prostitutes have been subject to selection bias because prostitutes were sampled at clinics, prisons, or treatment programs. The current study minimized such bias by sampling all street prostitutes in a circumscribed area known to be frequented by prostitutes.

The purpose of this study was to assess risk behaviors and to quantify the prevalence of HIV-1, HBV, HCV, syphilis, and HTLV I/II in street prostitutes and others soliciting sex.

Methods

In July and August of 1990 we contacted all persons who appeared to be soliciting clients for sexual activity in a geographically defined area known to be frequented by prostitutes. Subjects were offered $20 to be interviewed and provide an anonymous blood sample. A standard questionnaire was used and one investigator (S.R.T.) was always present. Test results and posttest counseling were made available through street outreach and the public sexually transmitted disease (STD) clinic.

HIV-1 antibodies were detected by enzyme immunoassay and confirmed by immunofluorescence assay. HBV infection was assessed by core antibody-IgG (HBcAb) and surface antigen (HBsAg). Antibodies to HCV and HTLV I/II were detected by enzyme immunoassay. The micro-hemagglutination assay was used for *Treponema pallidum* antibodies.

Individuals were classified as prostitutes if they reported having ever received money or drugs for sex. Condom use was assayed by reported use during any sexual activity in the previous month.

All analyses were carried out with the EGRET statistical package.[9] Odds ratios and confidence intervals for univariate and multivariate effects were estimated by logistic regression.[10]

Results

Of the 120 persons approached, 109 (91%) agreed to participate. All accepted voluntary HIV-1 testing and counseling. Participant characteristics are shown in Table 1. Although all participants were soliciting sex, only 73% reported doing so to receive money or drugs. All reported clients were male. One third of the male prostitutes behaved and were dressed as women at the time of interview. Most study participants expressed great concern about their health status, especially

Stephen R. Tabet is with the Department of Internal Medicine and School of Public Health and Community Medicine, University of Washington, Seattle, Washington. Darwin L. Palmer is with the Division of Infectious Diseases, Department of Internal Medicine, and William H. Wiese, Ronald E. Voorhees, and Dorothy R. Pathak are with the Department of Family, Community, and Emergency Medicine—all at the University of New Mexico School of Medicine, Albuquerque. Darwin L. Palmer is also with the Veterans Affairs Medical Center, Albuquerque, and Ronald E. Voorhees is also with the Office of Epidemiology, New Mexico Department of Health, Albuquerque.

Requests for reprints should be sent to William H. Wiese, MD, MPH, University of New Mexico, Family Practice Center, 2400 Tucker Avenue, NE, Albuquerque, NM 87131.

This paper was submitted to the Journal May 8, 1991, and accepted with revisions February 24, 1992.

TABLE 1—Distribution of Age, Ethnicity, and Other Variables for Male and Female Street Study Participants

	Male (n = 66)		Female (n = 43)	
	No.	%	No.	%
Age, y				
<25	38	57.6	13	30.2
25–29	17	25.8	11	25.6
≥30	11	16.7	18	41.9
Unknown	0	0.0	1	2.3
Ethnicity				
Anglo	21	31.8	12	27.9
Afro-American	5	7.6	11	25.6
Hispanic	29	43.9	16	37.2
Other	11	16.7	4	9.3
Condom use				
Always	28	42.4	25	58.1
Often	17	25.8	9	20.9
Sometimes/never	21	31.8	9	20.9
Number of sexual partners per night				
0–1	39	59.1	6	14.0
2–5	22	33.3	22	51.2
>5	5	7.6	15	34.9
Prostitution				
No	26	39.4	3	7.0
Yes	40	60.6	40	93.0
Intravenous drug use				
No	32	48.5	10	23.3
Yes	34	51.5	33	76.7
Share needles				
No	11	33.3	8	24.2
Yes	22	66.7	25	75.8

the risk of HIV infection, and responded enthusiastically to education.

Three persons had antibodies to HIV-1. Two were Hispanic homosexual males who reported engaging in receptive anal intercourse; one of these also reported intravenous (IV) drug use. Both had antibodies to HBV, but not to HCV. The third person who was HIV-positive was a non-Hispanic white female prostitute who used IV drugs and had antibodies to HBV and HCV.

Fifty-eight percent of the study participants had antibodies to either HBV or HCV; 42% were negative for both. Of all participants, 14% had antibodies to HBV only, 20% had antibodies to HCV only, and 25% had antibodies to both. Seven percent had detectable HBsAg. Individuals reporting IV drug use had HBV and HCV rates of 46% and 64%, respectively. Eight (8%) had antibodies to syphilis. Two persons had antibodies to HTLV I/II (but not to HIV-1).

Table 2 shows the univariate associations of various factors with hepatitis antibodies. Variables associated with both HBV and HCV included increased age, female gender, and IV drug use. Individuals reporting IV drug use were 45 times more likely to have HCV antibodies than were those reporting non IV drug use. Of the 19 study participants who were younger than 20 years of age, none showed infection with HBV; one had antibodies to HCV. Females were six times more likely than males to have HCV antibodies. HCV was associated with both prostitution and increased number of sexual partners per night. Seropositivity to HBV was related to reported condom use. Neither ethnicity nor sharing needles was significantly related to HBV or HCV. Receptive anal intercourse was not associated with HBV or HCV infection. To test the independence of significant univariate associations, we used a stepwise logistic regression model (Table 3). Increased age, IV drug use, and reported use of condoms were independent risk factors for antibodies to HBV. For HCV, IV drug use and female gender were independent risk factors.

The effect of female gender on HCV infection was evaluated separately for IV drug users and nonusers. The odds ratios were 6.9 (95% CI = 1.7, 27.6) and 4.4 (95% CI = 0.3, 79.7) for drug users and nonusers, respectively. Thus, female gender remained significant only for IV drug users.

Condom use was compared for those older and younger than 25 years of age; the positive association of condom use with HBV remained significant only in those older than 25 years.

Discussion

The HIV-1 seroprevalence rate in this population was surprisingly low (3%), especially when compared with the 27% seroprevalence rate in homosexual men attending the local public STD clinic (New Mexico Department of Health, Office of Epidemiology, unpublished data, 1991). The relatively high prevalence of HBV and HCV allowed us to analyze risk factors known to be associated with HIV-1 transmission.

IV drug use was strongly associated with HBV and HCV infection. However, we found no significant association between self-reports of sharing injection equipment and HBV or HCV infection. An analogous lack of association between reported use of shared equipment and HIV infection has been documented.[11,12] Responses to questions about shared injection equipment may be unreliable because of reluctance to admit this behavior and because of ambiguous definitions of "sharing."[3]

Unexpectedly, reported use of condoms was positively associated with HBV infection, as has been previously observed,[13] though this association was statistically significant only for those older than 25 years of age. We speculate that high rates of *current* condom use may be associated with *past* high-risk behavior or exposures.

That receptive anal intercourse was not associated with HCV infection in our study sample is not surprising. Most published surveys suggest that HCV is inefficiently transmitted sexually, compared with HBV and HIV.[6,14,15] The lack of association between receptive anal intercourse and HBV is unclear. Self-reports of sexual activity may be unreliable; also, infection may have occurred prior to the period for which the behavior was assessed.[6]

Age was an important risk factor for HBV, reflecting the cumulative effect of risk behaviors. Female gender remained a risk factor for HCV. Subjectively, it

TABLE 2—Seroprevalence Rates and Univariate Odds Ratios (ORs) for Street Study Participants

	Hepatitis B				Hepatitis C			
	No. Positive/Total	%	OR	95% CI	No. Positive/Total	%	OR	95% CI
Age, y								
<25	12/51	24	1.0		15/51	29	1.0	
25–29	13/28	46	3.2*	1.2, 8.7	15/28	54	3.5*	1.3, 9.5
≥30	14/29	48	4.4**	1.6, 12.5	16/29	55	4.3**	1.5, 12.0
Gender								
Male	20/66	30	1.0		19/66	29	1.0	
Female	19/43	44	2.5*	1.1, 5.8	26/43	60	6.3***	2.5, 15.6
Ethnicity								
Anglo	9/33	27	1.0		14/33	42	1.0	
Afro-American	5/16	31	1.2	0.3, 4.4	7/16	44	1.0	0.3, 3.5
Hispanic	21/45	47	2.2	0.9, 6.0	22/45	49	1.2	0.5, 3.0
Other	4/15	27	1.0	0.3, 4.3	2/15	13	0.2	0.04, 1.1
Condom use								
Always	23/53	43	1.0		23/53	43	1.0	
Often	10/26	38	0.7	0.3, 2.0	11/26	42	0.9	0.3, 2.4
Sometimes/never	6/30	20	0.3*	0.1, 0.8	11/30	37	0.6	0.2, 1.5
Number of sexual partners per night								
0–1	14/45	31	1.0		13/45	29	1.0	
2–5	15/44	34	1.4	0.6, 3.6	19/44	43	2.5*	1.0, 6.1
>5	10/20	50	2.8	0.9, 8.5	13/20	65	6.4**	1.9, 21.6
Prostitution								
No	9/29	31	1.0		7/29	24	1.0	
Yes	30/80	38	1.6	0.6, 4.0	38/80	48	3.5*	1.3, 9.3
Intravenous drug use								
No	8/42	19	1.0		2/42	5	1.0	
Yes	31/67	46	4.5**	1.6, 10.4	43/67	64	45.4***	9.9, 207.5
Share needles								
No	6/19	32	1.0		12/19	63	1.0	
Yes	24/47	51	2.7	0.8, 8.5	31/47	66	1.4	0.4, 4.7
Receptive anal intercourse								
No	26/66	39	1.0		29/66	44	1.0	
Yes	13/38	34	0.7	0.3, 1.6	15/38	39	0.7	0.3, 1.6

*$P < .05$; **$P < .005$; ***$P < .001$.

appeared that female IV drug users had more needle-sharing partners than did males.

The population we studied is characterized by alarming rates of behaviors associated with transmission of infectious diseases, as evidenced by the 58% seroprevalence of HBV, HCV, or both. The current low rate of HIV-1 infection in this group signals an opportunity for preventive intervention. Historically, prostitutes and IV drug users have been assumed to be nonresponsive to preventive efforts. In our study, however, most individuals were very responsive to information and appeared likely to attempt to protect themselves by using condoms and clean needles. The very high level of study participation (91%) of prostitutes who were approached in their work environment in-

TABLE 3—Adjusted Odds Ratios (ORs) for Street Study Participants

	Hepatitis B		Hepatitis C	
	OR	95% CI	OR	95% CI
Intravenous drug use				
No	1.0	...	1.0	...
Yes	2.9**	1.0, 8.2	45.9***	9.1, 230.5
Gender				
Male	...	...	1.0	...
Female	...	...	6.4***	1.9, 21.9
Age				
Per year	1.12*	1.03, 1.22	...	...
Condom use				
Always	1.0		...	...
Often	0.6	0.2, 1.9	...	...
Sometimes/never	0.3*	0.1, 0.9	...	...

*$P < .05$; **$P < .01$; ***$P < .001$.

dicates that street-based outreach programs may be an effective method for reaching this population. □

Acknowledgments

This study was supported in part by a grant from the Health of the Public Program of the Pew Charitable Trust and Rockefeller Foundation and a grant from the University of New Mexico School of Medicine Dean's Fund.

We would like to thank the following persons for their assistance and contributions to the success of this project: Sarah Allen, Steve Baca, Elizabeth Cordova, Patti Doherty, Nick Keller, Lori Lambert, Nili Lange, Anna Levy, Steve Lopez, Sandra McCollum, Bogdan Pathak, Amber Staley, Rick Steece, and the 109 street study participants. Laboratory studies were provided by the Laboratory Division of the New Mexico Department of Health. Condoms were donated for distribution by the Gay Men's Health Project.

References

1. Turner CF, Miller HG, Moses LE, eds. *AIDS, Sexual Behavior and Intravenous Drug Use*. Washington, DC: National Academy Press; 1989:152–153.
2. Centers for Disease Control. Antibody to human immunodeficiency virus in female prostitutes. *MMWR*. 1987;36:157–161.
3. Darrow WW, Deppe DA, Schaffer A, et al. Prostitution, intravenous drug use, and HIV-1 infection in the United States. In: Plant MA, ed. *AIDS, Drugs, and Prostitution*. London: Routledge; 1990:28.
4. Wallace JI, Beatrice S. Survey of street-walking prostitutes in New York City for anti-HIV-1 antibodies. In: the proceedings of the Fifth International Conference on AIDS; 1989; Montreal, Canada. Abstract A.564.
5. Hyams KC, Escamilla J, Romero RL, et al. Hepatitis B infection in a non-drug abusing prostitute population in Mexico. *Scand J Infect Dis*. 1990;22:527–531.
6. van den Hoek JAR, van Haastrecht HJA, Goudsmit J, de Wolf F, Coutinho RA. Prevalence, incidence, and risk factors for hepatitis C virus infection among drug users in Amsterdam. *J Infect Dis*. 1990;823–826.
7. Bell J, Batey RG, Farrell GC, et al. Hepatitis C in intravenous drug users. *Med J Aust*. 1990;153:274–276.
8. Khabbaz RF, Darrow WW, Hartley TM, et al. Seroprevalence and risk factors for HTLV-I/II infection among female prostitutes in the United States. *JAMA*. 1990; 263:60–64.
9. Epidemiologic Graphics, Estimation, and Testing Package (EGRET). Seattle, Wash: Statistics and Epidemiology Research Corporation; 1987.
10. Rothman KJ. *Modern Epidemiology*. Boston, Mass: Little, Brown & Co; 1986.
11. Williams ML. HIV seroprevalence among male IVDUs in Houston, Texas. *Am J Public Health*. 1990;80:1507–1509.
12. McCoy CB, Khoury E. Drug use and the risk of AIDS. *Am Behav Scientist*. 1990; 33:419–431.
13. Esteban JI, Viladomiu I, Gonzalez A, et al. Hepatitis C virus antibodies among risk groups in Spain. *Lancet*. 1989;2:294–297.
14. McCusker J, Koblin B. Lewis BF, Su[illegible] J. Demographic characteristics, ri[illegible] haviors, and HIV seroprevalence amon[illegible] intravenous drug users by site of contact: results from a community-wide HIV surveillance project. *Am J Public Health*. 1990;80:1062–1067.
15. Eyster ME, Alter HJ, Aledort LM, et al. Heterosexual co-transmission of Hepatitis C virus (HCV) and human immunodeficiency virus (HIV). *Ann Int Med*. 1991; 115:764–768.

Call for Abstracts for International Conference on Chronic Diseases and Changing Care Patterns

Organized by the Netherlands Society for Public Health and Science and the University of Amsterdam, an international conference on "Chronic Diseases and Changing Care Patterns in an Aging Society" will be held in Amsterdam, the Netherlands, June 9-11, 1993.

A broad range of multidisciplinary topics on both somatic and psychiatric chronic conditions will be addressed. In the plenary sessions, international experts will give their views on research and ongoing programs and their implications for care and policy. Keynote speakers include Dr. Trudi van den Bos, Dr. Jacob A. Brody, Dr. Alain Colvez, Dr. Robbert Huijsman, Dr. Louise J. Gunning-Schepers, Dr. Henk J. J. Leenen, Dr. John B. McKinlay, Dr. S. Jay Olshansky, Dr. Walter W. Rosser, and Dr. Lois M. Verbrugge.

For the parallel sessions, you are invited to submit abstracts on these issues: epidemiological, care, psychological and social, policy, economic, and ethical and legal.

Abstracts must be submitted before *October 31, 1992*. For further details and information on submission and format of abstracts, please write to: Dr. Trudi van den Bos or Wien Limburg, Institute of Social Medicine, University of Amsterdam, Meibergdreef 15, 1105 AZ Amsterdam, The Netherlands; fax 31 20.6912401.

Decreased Access to Medical Care for Girls in Punjab, India: The Roles of Age, Religion, and Distance

Beverley E. Booth, MD, MPH, and Manorama Verma, MD, DCH, MB, BS

ABSTRACT

Risk factors that increase the likelihood of discrimination against girls in India have not been well studied. In this study of hospitalized children in Punjab, India, girls were less likely to be in the newborn or infant age groups, to be of the Sikh religion, or to come from far away than were boys. These differences suggest that these factors are significant risk factors for denied access to medical care for girls living in Punjab, India. (*Am J Public Health.* 1992;82:1155–1157)

Introduction

Discrimination against girls in India, especially in the northern region, has been well documented.[1–6] Although lower socioeconomic status and high female sibling birth order have been associated with greater discrimination against girls,[2,6] other risk factors have been less well studied. In this paper we report on the effects of age, religion, and distance from the hospital on access to medical care in girls in Punjab, northwestern India.

Materials and Methods

Data were obtained on 2595 of the 2628 children admitted during 1987 to the Pediatric Ward of the Brown Memorial Hospital, Christian Medical College, Ludhiana, Punjab, India.

The male/female (M/F) ratio is the number of males admitted divided by the number of females admitted. Statistical analyses were performed using BMDP Statistical Software (Los Angeles, Calif) and Epi Info (Centers for Disease Control, Atlanta, Ga). The odds ratio is an estimate of the "odds," or likelihood, that a member of a particular group, compared with the reference group, will be female. Stepwise logistic regression was used to model the odds of being female, with simultaneous adjustment of three variables: (1) age (children [1 to 14 years of age], infants [1 to 12 months of age], and newborns [<1 month of age]), (2) religion (Hindu or Sikh), and (3) distance (coming from the city of Ludhiana and coming from outside the city of Ludhiana). The initial model contained all possible cross-product (interaction) terms, because for each combination, interaction was plausible because of prevailing social, economic, and cultural practices. The best-fit model was determined by elimination of nonsignificant ($P > .05$) terms, starting with the least significant interaction term, until all variables remaining were significant. The odds ratios and their 95% confidence intervals (CI) were calculated from the regression coefficients and covariance matrix of coefficients produced by the logistic model.

Results

Of the 2595 patients, 506 (19.5%) were girls, resulting in a M/F ratio of 4.13 (Table 1). The M/F ratio was highest in the newborn age group, in Sikhs, and in those coming from farthest away. The M/F ratio fell rapidly over the first few months of life, continued to decline through early childhood, and then leveled off after 4 years of age (Figure 1).

The unadjusted effects of age, religion, and distance are shown in Table 2. The likelihood of being female decreased with decreasing age and with increasing distance from the hospital, and was lower in Sikhs than in Hindus. The odds ratios determined by logistic regression are reported in Table 3 (Hindu children living in Ludhiana were set as the reference group). The Hosmer-Lemeshow goodness-of-fit χ^2 was 3.818 ($P = .873$). The final model consisted of age, religion, distance, and the interaction terms age × religion and distance × religion.

Discussion

The magnitude of restricted access to hospital care for girls shown by our data is impressive: the M/F ratio of 4.13 suggests that about three out of four girls (75%) who are ill enough to require hospitalization are denied this essential medical care simply because of their sex. Although it might be argued that the higher proportion of boys being admitted to the hospital resulted from a higher proportion of serious illness among boys, this is unlikely, be-

At the time of the study, the authors were with the Department of Pediatrics, Christian Medical College, Ludhiana, Punjab, India. Beverley E. Booth is now with the Department of Social and Preventive Medicine, Christian Medical College.

Requests for reprints should be sent to Beverley E. Booth, MD, MPH, Department of Social and Preventive Medicine, Christian Medical College, Ludhiana, Pb 141008, India.

This paper was submitted to the Journal June 25, 1991, and accepted with revisions February 25, 1992.

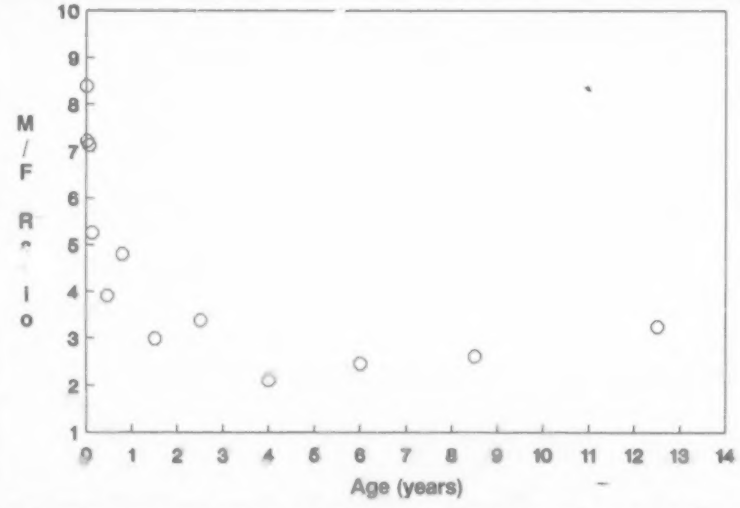
TABLE 1—Age, Religion, and Distance from Hospital for Children Admitted to Pediatric Ward, Christian Medical College, Ludhiana, 1987

	Girls[a] (n = 506) No. (%)	Boys[b] (n = 2089) No. (%)	Male/Female Ratio[c]
Age group			
Newborn (<1 month)	84 (16.6)	628 (24.2)	7.48
Infant (1 to 12 months)	158 (31.2)	731 (35.0)	4.63
Child (1 to 14 years)	262 (51.8)	726 (34.8)	2.77
Not known	2 (<0.1)	4 (<0.1)	4.00
Religion			
Sikh	188 (37.2)	1008 (48.3)	5.36
Hindu	305 (60.3)	1046 (50.1)	3.43
Other	13 (2.6)	35 (1.7)	2.69
Distance from hospital			
>80 km	52 (10.3)	403 (19.3)	7.75
60 to 80 km	44 (8.7)	285 (13.6)	6.48
<60 km	111 (21.9)	570 (27.3)	5.14
Ludhiana City	256 (50.6)	702 (33.6)	2.74
Not known/out of state	43 (8.5)	129 (6.2)	3.00

[a]Mean age was 30.1 months (SD = 37.6[a]); median age was 12 months.
[b]Mean age was 20.9 months (SD = 34.5); median age was 5 months.
[c]The overall male/female ratio was 4.13.
[d]Compared to boys, $P < .0001$ by Student's *t* test.

cause in Punjab the infant and child mortality rates are higher for girls than for boys.[7]

The lower rate of hospital admissions for girls than for boys is congruent with social factors that result in a strong preference for sons. The presence of a dowry system, a culture in which sons inherit family land and are responsible for taking care of elderly parents, and the lesser economic worth of women, who are less well educated and less likely to work outside the home than are men, all contribute to this preference.[1]



Note. Values represent the M/F ratio of children of the following ages: 1 day, 2 to 7 days, 8 to 31 days, 1 to 3 months, 4 to 6 months, 7 to 11 months, 1 year, 2 years, 3 to 4 years, 5 to six years, 7 to 9 years, 10 to 14 years. Each group was composed of at least 123 children, including at least 21 females.

FIGURE 1—Male/female (M/F) ratio by age in children admitted to Christian Medical College, Ludhiana, 1987.

The very high M/F ratio of 7.48 in newborns admitted to the hospital must reflect these intense and complex pressures. It is a reasonable inference that these pressures lead approximately six out of seven families in the catchment area of this hospital to deny hospital admission to newborn girls ill enough to require hospitalization. The rapid decline in the M/F ratio during the first few months of life suggests that the emotional bonding between an infant girl and her family in the first weeks of life might override the factors that prevent families from seeking appropriate medical care for newborn girls.

A profound effect of distance from the hospital on access to medical care for girls is also demonstrated in the present study. Distance from a health facility is a substantial barrier to receiving medical care for anyone, but this may be especially true for children who live in a state such as Punjab, in which civil strife often makes travel difficult, uncertain, and unsafe. In addition, cultural mores in Punjab dictate that women should not travel unaccompanied even in untroubled times. Hence a trip to the hospital may represent a substantial consumption of limited resources in terms of money and work time lost for adult male relatives who must accompany the child.

In this study, hospitalized girls were less likely to be Sikhs than were hospitalized boys; however, this finding was primarily true for children who were younger than 1 year of age and for those who came from outside Ludhiana. It is possible that the admission of fewer Sikh girls from outside Ludhiana, compared with Hindu girls from outside Ludhiana, might in fact be a function more of accessibility than of religion. In Punjab, Sikhs are more likely to live in villages, whereas Hindus are more likely to live in cities, where transportation systems are better.

This study demonstrates that for girls in Punjab, young age, increasing distance from the hospital, and Sikh religion decrease access to hospital care. To redress discrepancies in access to medical care, local health workers might be instructed to make more frequent visits to homes with girls in high-risk categories. Studies are under way to determine the role that emotional attachment of the caretaker to the infant might play in determining access to medical care for girls in Punjab. If attachment turns out to be an important factor, then health workers can be trained to detect infant girls who

TABLE 2—Effects of Age, Religion, and Distance from the Hospital on the Likelihood of Being Female in Children Admitted to Pediatric Ward, Christian Medical College, Ludhiana, 1987

	Odds Ratio	95% CI
Age group		
Newborn (<1 month)	0.37	0.28, 0.49
Infant (1 to 12 months)	0.60	0.48, 0.75
Child[a] (1 to 14 years)	1.00*	...
Religion		
Sikh	0.64	0.52, 0.79
Hindu[a]	1.00	...
Distance from hospital		
>80 km	0.35	0.25, 0.49
60 to 80 km	0.42	0.29, 0.62
<60 km	0.53	0.41, 0.69
Ludhiana City[a]	1.00**	...

[a]Reference group.
*Linear test for trend, $\chi^2 = 58.94$, $P < .00001$.
**Linear test for trend, $\chi^2 = 55.00$, $P < .00001$.

TABLE 3—Effects of Age, Religion, and Distance from Hospital on the Likelihood of Being Female in Children Admitted to Pediatric Ward, Christian Medical College, Ludhiana, 1987

	Children (1 to 14 years)	Infants (1 to 12 months)	Newborns (<1 month)
From Ludhiana			
Hindu	1.00[a]	0.81 (0.61, 1.09)	0.52 (0.36, 0.76)
Sikh	1.51 (1.04, 2.21)	0.64 (0.43, 0.97)	0.51 (0.30, 0.85)
From outside Ludhiana			
Hindu	0.74 (0.56, 0.97)	0.59 (0.37, 0.96)	0.38 (0.26, 0.58)
Sikh	0.38 (0.27, 0.54)	0.24 (0.16, 0.35)	0.19 (0.12, 0.29)

Note. The numbers shown are odds ratios for being female, modeled by logistic regression; 95% CIs are shown in parentheses.
[a]Reference group.

are at risk for this reason and to enhance the attachment process. □

Acknowledgments

This research was partially supported by a grant from the Population Sciences Section of the Rockefeller Foundation and by salary support of Dr. Booth by the Presbyterian Church (USA).

Data from this study were presented at a national conference, Status of the Girl Child in India, sponsored by the Indian Academy of Pediatrics on December 8, 1990, in Bombay, India.

References

1. Miller BD. *The Endangered Sex: Neglect of Female Children in Rural North India.* Ithaca, NY: Cornell Univeristy Press; 1981.
2. Gupta MD. Selective discrimination against female children in rural Punjab, India. *Popul Dev Rev.* 1987;13:77–100.
3. Mandelbaum DG. *Society in India.* Berkeley, Calif: University of California Press; 1970.
4. Khosla T. The plight of female infants in India. *J Epidemiol Community Health.* 1980; 34:143–146.
5. Taylor CE, Kielmann AA, DeSweemer C, et al. The Narangwal experiment on interaction of nutrition and infections: I. project designs and effects upon growth. *Indian J Med Res.* 1978;68s:1–20.
6. Cowan B, Dhanoa J. The prevention of toddler malnutrition by home-based nutrition health education. In: McLaren DS, ed. *Nutrition in the Community.* Chichester, England: John Wiley & Sons; 1983:339–356.
7. Holla M. Vital statistics system—a major source of information on infant and child mortality. *Indian J Pediatr.* 1985;52:115–126.

A8623-1

American Journal of Public Health

September 1992, Vol. 82, No. 9 Established 1911

86389 028
ED KIANKA
CONGRESSIONAL INFORMATION SVC
4520 EAST-WEST HIGHWAY
BETHESDA MD 20814

APHA
American Public Health Association

Featuring Smoking and Health

Editorial

1203 The Vital Diversity of Tobacco Control Research *G. A. Giovino, M. P. Eriksen, J. W. McKenna*

1991 Presidential Address

1206 Creating a Healthy Public *C. W. Keck*

Featuring Smoking and Health

1210 Communitywide Smoking Prevention: Long-Term Outcomes of the Minnesota Heart Health Program and the Class of 1989 Study *C. L. Perry, S. H. Kelder, D. M. Murray, and K.-I. Klepp*

1217 Locking Devices on Cigarette Vending Machines: Evaluation of a City Ordinance *J. L. Forster, M. E. Hourigan, and S. Kelder*

1220 Adolescents' Multisubstance Use Patterns: The Role of Heavy Alcohol and Cigarette Use *S. L. Bailey*

1225 Determinants of Passive Smoking in Children in Edinburgh, Scotland *M. J. Jarvis, D. P. Strachan, and C. Feyerabend*

1230 Cigarette Smoking and Occupational Status: 1977 to 1990 *L. S. Covey, E. A. Zang, and E. L. Wynder*

1235 Predictors of Smoking Relapse after Noncardiac Surgery *J. A. Simon, W. S. Browner, and D. T. Mangano for the Study of Perioperative Ischemia (SPI) Research Group*

1238 Smoking Cessation in Women Concerned about Weight *P. L. Pirie, C. M. McBride, W. Hellerstedt, R. W. Jeffery, D. Hatsukami, S. Allen, and H. Lando*

1268 Predictors of Smoking Cessation and Relapse in Older Adults *M. E. Salive, J. Cornoni-Huntley, A. Z. LaCroix, A. M. Ostfeld, R. B. Wallace, and C. H. Hennekens*

1271 The Tobacco Institute's "It's the Law" Campaign: Has It Halted Illegal Sales of Tobacco to Children? *J. R. DiFranza and L. J. Brown*

(continued page 1193)

Contents

American Journal of Public Health

The Journal's one-page guidelines for authors, "What *AJPH* Authors Should Know," are printed in each issue. They can also be obtained from the Editorial Office.

September 1992, Vol. 82, No. 9 ISSN: 0090-0036 Established 1911

Articles

1244 How Women's Adopted Low-Fat Diets Affect Their Husbands *A. L. Shattuck, E. White, and A. R. Kristal*

1251 Weight Loss Attempts in Adults: Goals, Duration, and Rate of Weight Loss *D. F. Williamson, M. K. Serdula, R. F. Anda, A. Levy, and T. Byers*

1258 Intestinal Parasites among North Carolina Migrant Farmworkers *S. D. Ciesielski, J. R. Seed, J. C. Ortiz, and J. Metts*

1263 Restricted Activity Days among Older Adults *M. R. Kosorok, G. S. Omenn, P. Diehr, T. D. Koepsell, and D. L. Patrick*

Public Health Briefs

1273 PHLIS: An Electronic System for Reporting Public Health Data from Remote Sites *N. Hargrett Bean, S. Morris Martin, and H. Bradford, Jr*

1277 Homosexual and Bisexual Men's Perceptions of Discrimination in Health Services *N. E. Kass, R. R. Faden, R. Fox, and J. Dudley*

Public Health Then and Now

1280 Making Human Junk: Child Labor as a Health Issue in the Progressive Era *A. Derickson*

Other Journal Departments

1194 APHA Masthead

1198 Advertisers' Index

1200 What *AJPH* Authors Should Know

1203 Journal Masthead

1291 Letters to the Editor

1298 Book Corner

1303 Annual Meeting Registration Form and Information

1307 Job Opportunities

APHA
American Public Health Association

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.

APHA OFFICERS

Joyce C. Lashof, MD, *President*

C. William Keck, MD, MPH, *Immediate Past President*

Helen Rodriguez-Trias, MD, *President-Elect*

Arnold E. Greenberg, SM, *Vice-President (USA)*

Trevor Hancock, MB, MHSc, *Vice-President (Canada)*

Julio Teja Perez, MD, *Vice-President (Latin America)*

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Frances M. Veverka, RS, MPH, *Chair, Executive Board*

George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*

William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD

Frances M. Veverka, RS, MPH (1992) *Chair*

Marilyn Aguirre-Molina, EdD, MS (1993)

Brian A. Dobrow, MPH (1992)

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Kristine M. Gebbie, BSN, MN (1994) *Vice-Chair*

Alan R. Hinman, MD, MPH (1995)

Janet Heinrich, DrPH, MPH (1993)

Margaret H. Jordan, RN, MPH (1992)

C. William Keck, MD, MPH, *Immediate Past President*

Margo D. Kerrigan, MPH (1995)

Joyce C. Lashof, MD, *President*

Barry S. Levy, MD, MPH (1993)

Helen Rodriguez-Trias, MD, *President-Elect*

Allan G. Rosenfield, MD (1994)

Reed V. Tuckson, MD (1994)

Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO

William H. McBeath, MD, MPH, *Executive Director*

Joanne E. Lukomnik, MD, *Chair, Action Board*

George Pickett, MD, MPH, *Chair, Program Development Board*

George E. Hardy, MD, MPH, *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc, 1015 Fifteenth St, NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50 000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health,* 1015 Fifteenth St, NW, Washington, DC 20005.

Information on Manuscript Preparation is published monthly in the Journal. The Journal publishes reports of original research, research methods, program evaluations, analytic reviews, including health policy analysis, and reports for special departments. A paper is considered with the understanding that it or its essential substance neither has been published nor is under consideration elsewhere.

Subscription price, individual, is $100/year US, $140/year international; **institutional,** $160/year US, $200/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to *American Journal of Public Health,* 1015 Fifteenth St, NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St, NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St, NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS/COLLEAGUE, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc, Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Communitywide Smoking Prevention: Long-Term Outcomes of the Minnesota Heart Health Program and the Class of 1989 Study

Cheryl L. Perry, PhD, Steven H. Kelder, MPH, David M. Murray, PhD, and Knut-Inge Klepp, PhD

ABSTRACT

Objectives. The Class of 1989 Study is part of the Minnesota Heart Health Program (MHHP), a populationwide research and demonstration project designed to reduce cardiovascular disease in three educated communities from 1980 to 1993. This paper describes an intensive, school-based behavioral intervention on cigarette smoking, comparing long-term outcomes in one of the intervention communities with those in a matched reference community.

Methods. Beginning in sixth grade (1983), seven annual waves of cohort and cross-sectional behavioral measurements were taken from one MHHP intervention community and its matched pair. All students in each community were eligible to participate (baseline n = 2401). Self-reported data collected at each period described prevalence and intensity of cigarette smoking.

Results. There were no differences at baseline for either weekly smoking prevalence or intensity of smoking. Throughout the follow-up period, however, smoking rates as determined by these measures were significantly lower in the intervention community: 14.6% of students were weekly smokers at the end of high school compared with 24.1% in the reference community.

Conclusions. These results suggest that multiple intervention components such as behavioral education in schools, booster programs to sustain training, and complementary communitywide strategies may all be needed for lasting reductions in adolescent tobacco use. (*Am J Public Health.* 1992;82:1210–1216)

Introduction

Despite repeated disease warnings and some decline in the prevalence of cigarette smoking over the past 15 years, about 29.4% of the 1990 American high school graduates are current smokers.[1] They have acquired a pernicious habit that is difficult to give up and that has been described as the No. 1 cause of premature morbidity and mortality in the United States.[2] The peak time of smoking onset occurs in early adolescence, sixth to ninth grades,[3] when behavioral choices are complicated by biological, emotional, cognitive, identity, and social changes. The initiation of smoking, then, is embedded in these changes, which should be monitored and guided by the adult society.

Efforts to prevent the onset of cigarette smoking have been targeted primarily at the individual adolescent through school-based programs. These programs teach young adolescents the skills they need both to resist the social influences to smoke and to perform more general social competencies.[4,5] However, while the results of more than 20 research studies permit optimism that these approaches can delay onset of cigarette smoking, none has demonstrated positive outcomes after 3 years. Recently, several investigators have begun to add intervention components that explicitly target the social environment to aid adolescents in remaining nonsmokers.[6–9]

The Class of 1989 Study was designed to test the efficacy of a smoking prevention program that is part of a larger effort to reduce heart disease in whole communities. It is a substudy of the Minnesota Heart Health Program (MHHP), a populationwide, community-based cardiovascular disease prevention program funded by the National Heart, Lung, and Blood Institute over a 13-year period (1980 to 1993).[10] In the MHHP, the entire population in three participating cities in the north-central United States received a 5-year educational program encouraging healthy changes in eating, exercise, and smoking patterns to control high blood pressure. Three similar cities, also in the north-central portion of the United States, served as reference communities.[11] The Class of 1989 Study evaluated the youth education portion in two MHHP communities.

It was hypothesized that the effects of a school-based smoking prevention program (the Minnesota Smoking Prevention Program) with young adolescents would more likely be maintained if the program was part of 5 years of behavioral health programs in the schools and if these programs were implemented in communities where (1) adults were involved in risk factor screening and mass smoking cessation, (2) new smoking ordinances at school and in the community were being considered, and (3) multiple complementary school and community programs were established.

Methods

Research Design

The research design of the Class of 1989 Study was determined largely by the

Cheryl L. Perry, Steven H. Kelder, and David M. Murray are with the School of Public Health at the University of Minnesota in Minneapolis; Knut-Inge Klepp is with the University of Bergen in Norway.

Requests for reprints should be sent to Cheryl L. Perry, PhD, Division of Epidemiology, School of Public Health, University of Minnesota, 1300 S Second St, Suite 300, Minneapolis, MN 55454-1015.

This paper was submitted to the Journal January 27, 1992, and accepted with revisions June 8, 1992.

Editor's Note. See related editorial by Giovino et al. (p 1203) in this issue.

parent project, the MHHP. Two of the six MHHP communities were involved. The community made up of Fargo and West Fargo, North Dakota, and of Moorhead, Minnesota, was one of the three MHHP intervention communities, and it received an extensive communitywide intervention program designed to improve eating, exercise, and smoking patterns across the entire population. The youth smoking prevention program was one component of that communitywide effort. Sioux Falls, South Dakota, was the reference community matched to the Fargo-Moorhead intervention community for size, socioeconomic makeup, and distance from Minneapolis–St. Paul, Minnesota. There were no MHHP-sponsored education programs for either youth or adults in Sioux Falls.

The Class of 1989 Study was established on a quasi-experimental[12] basis to assess the net effect of combining a behavioral health education program in school with the communitywide MHHP on the youth living in the intervention community. All sixth graders enrolled in the public schools in both communities were invited to participate in a baseline survey in April 1983, and that grade cohort was surveyed annually each April until its graduation from high school in 1989. From 1983 to 1986, 13 grade schools participated in the survey; and from 1987 to 1989, participating students attended 7 high schools. All students in the grade cohort were eligible to participate in each annual survey regardless of their prior participation in either the Class of 1989 surveys or the MHHP education programs, or of how long they had resided in the community. Thus, each annual survey represented a cross-sectional sample of all students in each community for the appropriate grade. By recording identifying information at each survey, it was possible to analyze the data both as specific for a cohort and as representative of a series of cross-sectional surveys using all respondents. Cohort data, using only those respondents present at baseline and at subsequent surveys, provide the strongest basis with which to test the continuing effects of the interventions. While cross-sectional data may better reflect the secular trend, they also represent students new to the area who have not received the full treatment.

During this same period, the students in the Class of 1989 study intervention community participated in MHHP-sponsored behavioral health programs in school from 1983 to 1987. Smoking prevention programs, based on the social influences model, were implemented in 1983, 1984, and 1985. Details on all the interventions of the Class of 1989 Study and of the social influences model are described in detail elsewhere.[6,13]

The primary smoking intervention, the Minnesota Smoking Prevention Program, was implemented in 1984, the fall of seventh grade. This program addresses the prevention of tobacco use by influencing the social and psychological factors that encourage the onset of smoking. First, the students begin by identifying the short-term consequences of smoking, such as smelling like smoke or having bad breath. This generally is done through small group discussions so that the consequences are relevant to the age group. Second, the students discover that smoking is not a normative behavior for young adolescents; they do this by comparing their expectations of how many of their peers smoke with actual data and by discussing their overestimates of prevalence. Third, the reasons why adolescents smoke tobacco are explored—for example, a desire to have fun, a way of making friends, a signal of maturity, or a method of coping with personal problems[14]—and positive alternatives to achieve the same goals are provided. Fourth, the students learn how these associations are established in our culture through both advertising and peer and adult role models. The methods advertisers use to convince adolescents of tobacco's functional values are presented through discussions of selected advertisements. Mock social situations are analyzed to identify the type of influences that exist. Fifth and most important, the students learn and practice skills to resist these social influences to smoke. They create anti-tobacco advertisements and skits (role playing) around possible social encounters. Sixth, near the end of most programs, the students make a public commitment to abstain. This commitment acts as a psychological anchor and explicitly creates an intention not to experiment with substances. Finally, all these activities are experiential—designed to require active participation—and often are led by trained peer (same-age) leaders, who have proven to be effective communicators for many of the social and psychological messages intrinsic to the program.[15]

In addition to the school-based intervention, the intervention community participated in the MHHP populationwide intervention. Indirectly, the students received this intervention over a 5-year period through a deliberate attempt to restructure the adult social and physical environment with regard to heart-healthy behaviors and values.[10,11] The primary intervention components of the MHHP included seven strategies: (1) population-based risk factor screening for cardiovascular disease, in which more than 60% of adults participated in a 1-hour health assessment and direct education program for healthy eating, exercise, and smoking; (2) grocery and restaurant heart-health point-of-purchase food labeling education; (3) community organization and citizen task forces to develop annual risk factor education campaigns; (4) continuing education of health professionals to promote understanding of the nature of cardiovascular disease risk and prevention; (5) mass media education via television, radio, newsprint, etc.; (6) adult education in work sites, churches, workshops, and other organizations; and (7) youth education.

Outcome Evaluation

Outcome evaluation was conducted through an annual survey in all the school districts in the two communities from March to April of 1983 to 1989. The survey instrument was derived from a cognitive-behavioral model and the needs assessment surveys in each community.[13,16] Administered by survey staff from the MHHP who were trained for classroom administration, the survey took approximately 40 minutes and was completed by the students in their health, social science, or English classes. All students in the relevant grade level who were present on a particular day were tested.

The measures for tobacco use—the focus of this paper—each year included self-reported smoking history and, in 1986 only, the amount of thiocyanate in saliva samples. Psychosocial variables pertaining to smoking etiology were not included in this study because this question was the subject of a separate, ongoing research project and limited space was available on the survey instrument. The self-report items, standardized questions with acceptable reliability and validity, have been used in previous smoking research.[17] The prevalence of weekly smoking was used to assess the impact of the intervention. In addition, a smoking intensity score was created from the self-reported measures. Smoking intensity represents the number of cigarettes smoked per week by each student, calculated as an average of three self-report questions related to daily smoking, weekly smoking, and smoking history. The method of calculation of

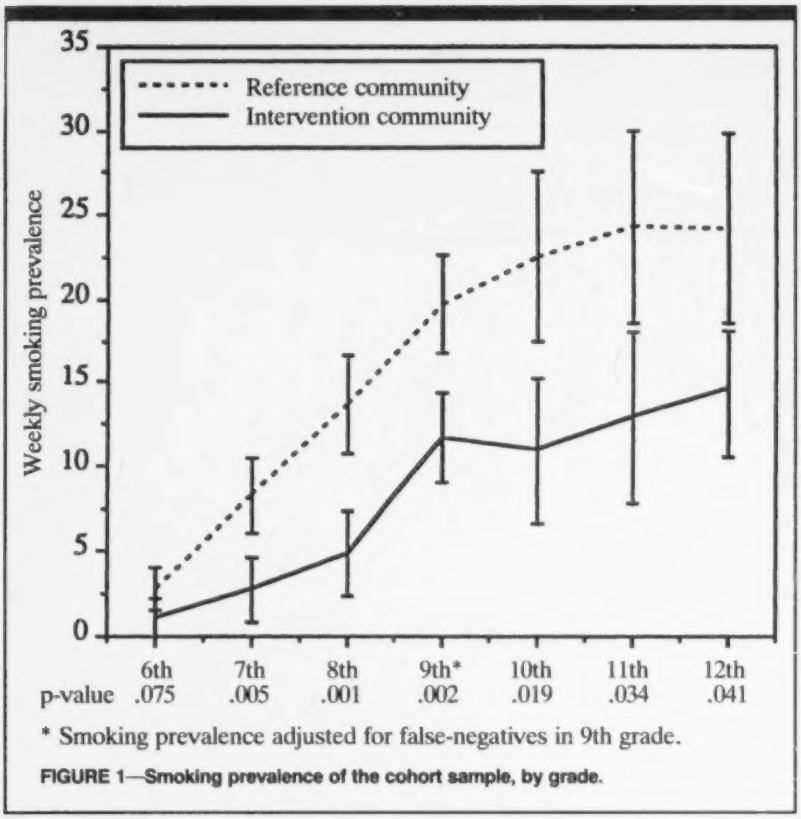


FIGURE 1—Smoking prevalence of the cohort sample, by grade.

smoking intensity has been described in detail elsewhere.[17] Test-retest correlation for smoking intensity from this age population was .99.[17]

Saliva samples were obtained for biologic validation of self-report in 1986 with a random sample of half the classrooms (n = 1076 students); this was done with dental cotton rolls, which were held in the mouth for 5 minutes. The amount of saliva thiocyanate (SCN), a byproduct of smoking tobacco, was collected and determined by methods described elsewhere.[17,18] The correlation of smoking intensity with SCN was 0.46 ($P < .0001$), a value not unexpected considering that student smoking habits at this age are sporadic. SCN is a useful measure in this situation primarily to identify false-negative self-reports (i.e., self-reported nonsmokers with an SCN value indicating smoking status) and is less likely to classify false-positive self-reports (self-reported smokers with a low SCN value) correctly. SCN values were used in these data to correct false-negative self-reports in 1986 using a cutoff point of 79 μg/mL derived from internal data as described by Cummings and Richard[19]; this accounts for low levels of SCN due to passive smoking.

Other relevant measures are age at the time of survey, sex, and parental occupation. Questions related to parental occupation were added in 1987 and are used to produce an index of household occupation in which 1 = both parents are white-collar workers, 2 = both parents are clerical workers, 3 = both parents are blue-collar workers, and 4 = parents are mixed as to the three categories. Single parents were coded for their occupation.

Analysis Methods

This study combines elements of what we have termed *nested cohort* and *nested cross-sectional* designs.[20] In these designs, intact social groups (e.g., whole cities) are assigned to conditions. In the nested cohort design, subjects within those groups are followed over time as a cohort to assess the effects of the interventions. In the nested cross-sectional design, independent samples of subjects within those social groups participate at each measurement occasion. In either design, subjects from a single community will tend to be more like one another than they are like subjects in other communities. This within-community correlation in the data, indexed by the community-level intraclass correlation coefficient (ICC), adds an additional component of variability to the intervention group means above and beyond that attributable to either the individual subjects or the interventions themselves.[21] Unless this component is accounted for in the analysis, the evaluation of the intervention effects will be positively biased in proportion to the magnitude of the ICC and the number of respondents in each community.[22]

Note that in the present study, students were observed in classrooms, which were nested within schools, which were nested within the towns assigned to the two treatment conditions. Fortunately, ICC, which reflects the within-unit correlation in the data at each level of nesting, also captures the within-unit correlation at all levels of nesting lower in the hierarchy. Thus, while we could not use the community as the unit of analysis, we could correct for most of the variance inflation expected in such a design by using instead the next level in the hierarchy: the school. This was accomplished for the data gathered after 1985 using PROC GLM in SAS,[23] in which a model that used the individual's smoking status as the outcome and included school as a nested random effect was fit for each round of data; the intervention effect was tested against the school variance. The school identifier was not available in the data collected for elementary schools prior to 1986, and so the school variance was estimated by standard methods using schools from subsequent years.[20,21] Separate analyses were performed for each follow-up round. For the cross-sectional analyses, all participants were included; for the cohort analyses, only those who also participated in the baseline survey were included. All analyses included covariance adjustments for age and sex, and in the years 1987 to 1989, adjustments were made for parental job class.

Results

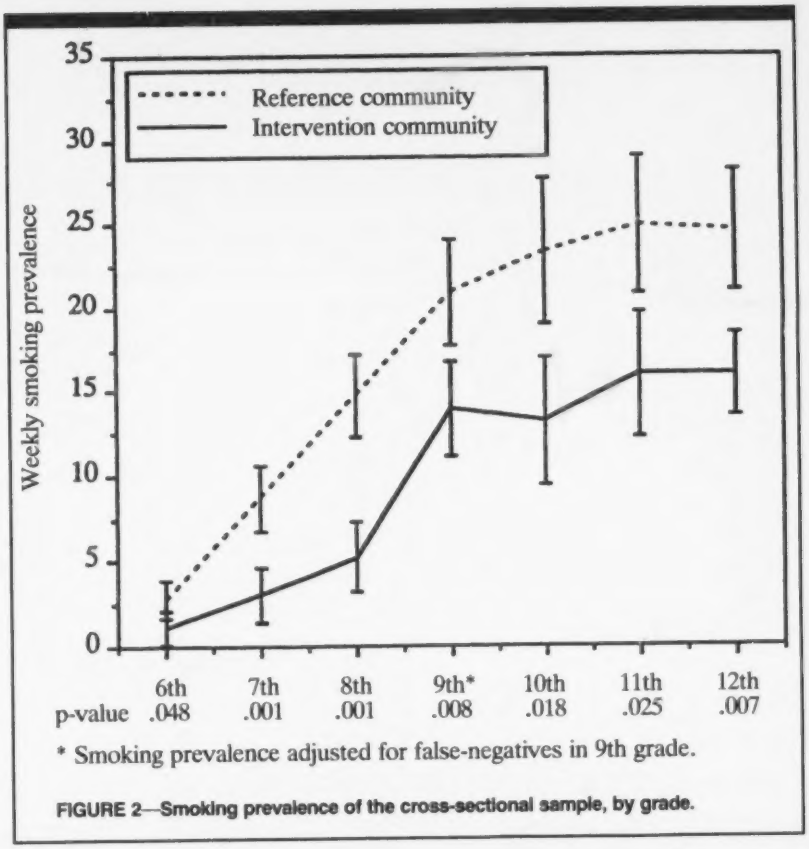
Smoking Outcomes

The cross-sectional and cohort smoking outcome data were analyzed to examine differences between communities in both the percentage of weekly smokers and the smoking intensity for each year from 1983 to 1989. No significant differential program effects were observed between the sexes, so the results for the two sexes were pooled for presentation purposes. Figures 1 and 3 present the results from the cohort analyses for

weekly smoking prevalence and smoking intensity, respectively; Figures 2 and 4 present the results from the serial cross-sectional analyses. All figures present mean levels, 95% confidence intervals, and intervention/reference comparison *P* values, adjusted for the school effects as described in the Analysis section above. The unadjusted smoking prevalence and intensity differed by less than 1%, or by one cigarette per week, respectively, for either treatment condition, and these measures offer the same magnitude of intervention effects as the covariate adjusted results.

Although no significant differences were detected at baseline (1983) in the cohort sample, Figures 1 and 3 reveal significant differences for smoking prevalence and intensity between communities for this sample at all follow-up years (1984 to 1989). At the end of high school, 14.6% of the cohort sample from the intervention community were smoking compared with 24.1% from the reference community. Similarly, Figures 2 and 4 show significant differences between communities for the cross-sectional samples at all follow-up years. Cross-sectional students represent those remaining in the community throughout the study as well as those not present at baseline or new to the community. In Figures 2 and 4 we observe only a slight increase in smoking prevalence and intensity for the cross-sectional sample as compared with the cohort sample, indicating that these students have little impact on how these trends are interpreted.

Saliva samples were obtained from a random sample of approximately half of all classrooms in 1986, with SCN analyses completed on 667 and 409 students in the intervention and reference communities, respectively. Using school as the unit of analysis, significant differences between communities were found in the SCN level, with the intervention community significantly lower than the reference community in both the cohort sample (54 vs 39 μg/mL, $P = .0002$) and the cross-sectional sample (56 vs 41 μg/mL, $P = .0009$). Note that adjustment for false-negative self-reported smoking prevalence was done in 1986 as described in the Methods section above. The number of false-negative reports verified by SCN differed by community, with 9.8% (n = 40) in the reference community claiming falsely not to be smoking compared with 6.5% (n = 43) in the treatment community ($\chi^2 = 4.3$; $P = .038$; individual unit of analysis).[18] The effect of false-negative adjustment was to raise the prev-



* Smoking prevalence adjusted for false-negatives in 9th grade.

FIGURE 2—Smoking prevalence of the cross-sectional sample, by grade.

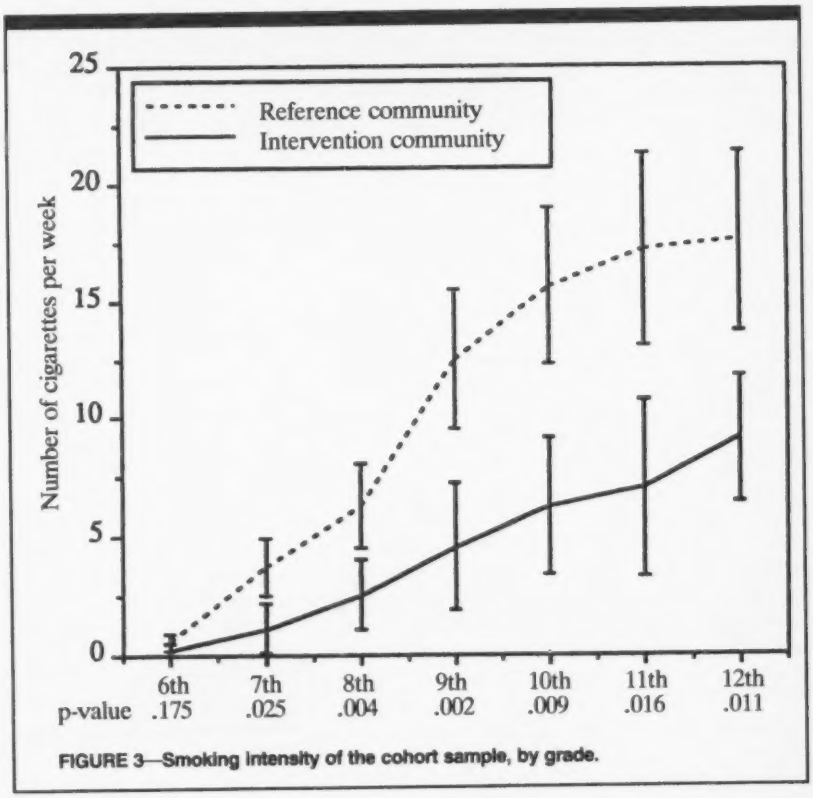


FIGURE 3—Smoking intensity of the cohort sample, by grade.

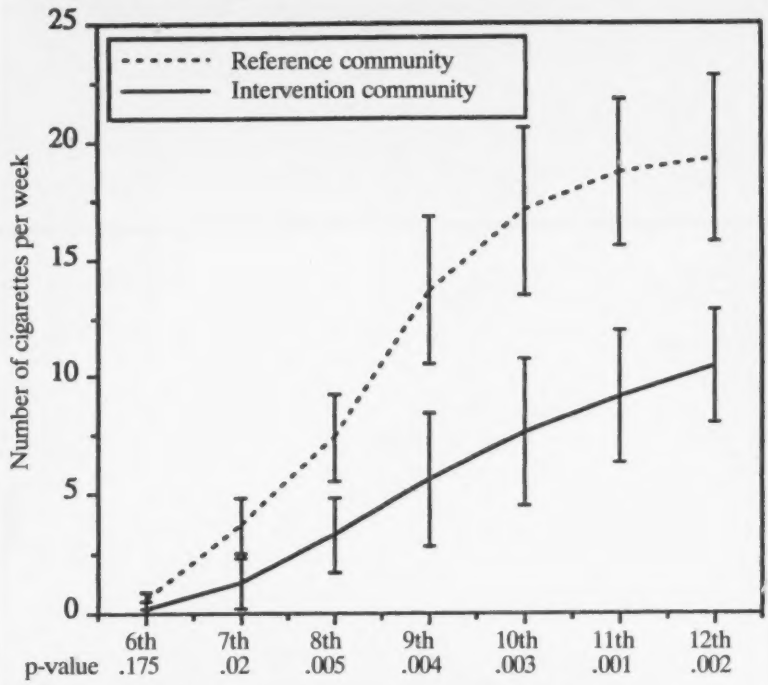


FIGURE 4—Smoking intensity of the cross-sectional sample, by grade.

TABLE 1—Sample Size and Follow-up by Year and Grade

	1983/ 6th	1984/ 7th	1985/ 8th	1986/ 9th	1987/ 10th	1988/ 11th	1989/ 12th
Cross-section, n	2401	2453	2416	2164	2124	1905	1439
Cohort, n	2401	2103	1943	1670	1578	1421	1080
Cohort, % follow-up	...	88	81	70	66	59	45
Cohort as % of cross-section	100	86	80	77	74	75	75

alence estimates for both communities. This is especially noticeable in Figures 1 and 2 for the intervention community, in which a sharp increase was observed between eighth and ninth grades, most of which can be attributed to the false-negative correction.

Attrition Analysis

Table 1 presents the study sample size and follow-up by year and grade for the cross-section and cohort samples. Financial constraints precluded any attempt to contact those students who moved out of the community, were absent, were involved in work-study, or dropped out and were not in class on the day of the survey. Response rates by intervention and reference communities were nearly equal except for the last year of data collection, in which 55% of intervention cohort students were resurveyed compared with only 32% in the reference schools. Because missing students were not recontacted, we cannot estimate the extent of differential attrition on smoking, but this can be approximated by identifying the previous year's values for those students not present. Table 2 presents the previous year's smoking prevalence for students missing in the cohort sample. For example, for students not present in 12th grade, 29% in the reference community were smoking at least weekly in 11th grade compared with 23% in the intervention community. Had those identified in any previous year remained smoking in the following year, the magnitude of the intervention difference would have been increased in four of the six follow-up evaluations by 0.5% to 2.3% and would have declined in the other two by 0.2% to 0.35%. In no year would the addition of these missing students have altered the intervention group differences observed or the given interpretation of these results.

Discussion

The Class of 1989 Study was designed to take advantage of the time of transition to adolescence by using peer-led, behavioral, school-based smoking prevention programs nested within a community-based intervention. This paper describes smoking outcomes for cross-sectional and cohort samples from the intervention and reference communities for 1983 to 1989.

Smoking rates throughout the follow-up period were substantially lower in the intervention community when measured by either weekly smoking prevalence or smoking intensity. These differences were confirmed by SCN analysis in 1986. The risk of being a smoker upon graduation in 1989 in the intervention community was about 40% lower than it was in the reference community.

The results of the study should be considered in comparison with other smoking prevention studies. For the most part, the effects of other programs have diminished over time, and few investigators have reported effects past the 10th grade.[9,24,25] This suggests that behavioral education in schools, booster programs to sustain training, and complementary communitywide change may all be needed to maintain effects with young people. Since virtually all cigarette smoking is initiated prior to high school graduation, the maintenance of significant differences through 12th grade should lead to the eventual reduction of adult smoking and smoking-related diseases.

The study has several limitations. The common threats to internal and external validity associated with nonrandom assignment are well known, and several such threats pertain to this design, including baseline pretreatment differences, regression to the mean, differential maturation and attrition, and history.[12] At baseline, both smoking prevalence and intensity measures were slightly higher in the reference community. These differences at the early stages of smoking onset were small, and separate analyses of covariance in the cohort sample adjusting for these differences required no changes in

TABLE 2—Previous Year's Smoking Prevalence of Students Missing in the Cohort Sample

	Identified Missing Students, n		Weekly Smoking, %	
	Intervention	Reference	Intervention	Reference
11th grade smoking prevalence of students missing in 12th grade	211	305	22.6	28.9
10th grade smoking prevalence of students missing in 11th grade	211	124	17.8	42.4
9th grade smoking prevalence of students missing in 10th grade	147	157	23.1	31.2
8th grade smoking prevalence of students missing in 9th grade	256	148	9.4	23.2
7th grade smoking prevalence of students missing in 8th grade	134	157	3.7	15.8
6th grade smoking prevalence of students missing in 7th grade	163	135	0.0	3.0

our interpretation of these data. Differential regression is unlikely, given the multiple observations and the fact that differences were observed after baseline observation and became larger over the 7 years. Attrition over time may underestimate smoking prevalence in both communities since absent students are more likely to smoke.[26,27] Attrition analysis indicates, for most years, a bias that reduced the magnitude of the intervention effect because of higher levels of smoking in the reference community as a result of those missing at the time of measurement.

Unfortunately, differential maturation and local history are completely confounded with the series of interventions, and thus they cannot be so easily dismissed. At issue is whether the lower onset rate in the intervention community was due to some cause other than the intervention, and whether the increase in smoking in the reference community was as high as it was due to some cause other than the usual high school secular trend. Annual reviews of related community programs as well as of the existing school health programs that were done by the investigators in either community did not reveal an alternative explanation for the effects observed. In addition, smoking rates at the completion of the study in Sioux Falls are comparable to those obtained in a 1990 representative survey of high school seniors for the North Central Region, in which 22.2% reported smoking at least one cigarette per day over the previous 30 days.[1] Finally, no policy level changes in taxes or availability of cigarettes were noted to have occurred during or immediately prior to the evaluation period.

The Class of 1989 Study does, then, permit optimism for the benefits to youth of long-term, school-based, communitywide programs, and specifically for the retardation of uptake of smoking behavior. The adoption of behavioral programs across multiple communities appears to be feasible; outcomes can be sustained in a supportive environment. The communities selected for this study represent typical middle-size midwestern communities, and the results should be generalizable to similar settings. Examination of these school-based, communitywide approaches appears to be warranted for youth with other high-risk social and health behaviors, and especially for youth in lower socioeconomic communities, where these problems are more acute and existing resources are scarcer. The need for research in these populations is clearly critical. □

Acknowledgments

This research was supported by the National Heart, Lung, and Blood Institute, Grant R01 HL 25523.

The authors wish to express thanks to Cynthia Sillers and Lori Steidl for their assistance in implementing and evaluating the Class of 1989 Study.

References

1. Johnston LD, O'Mally PM, Bachman JG. *Drug Use among American High School Seniors, College Students, and Young Adults, 1975–1990, Volume 1: High School Seniors.* Rockville, Md: National Institute on Drug Abuse; 1991.
2. Schultz, JM. Smoking-attributable mortality and years of potential life lost: United States, 1988. *MMWR.* 1991;40:62–71.
3. Johnson LD, O'Malley PM, Bachman JG. *Illicit Drug Use, Smoking, and Drinking by America's High School Students, College Students, and Young Adults.* Rockville, Md: National Institute on Drug Abuse; 1988.
4. Kelder SH, Perry CL. Substance abuse prevention. In: Parcel GS, Wallace HM, Patrick KP, Igoe J, eds. *Principles and Practices for School Health.* Oakland, Calif: Third Party Publishing (in press), 1992.
5. Flay, BR. Psychosocial approaches to smoking prevention: a review of findings. *Health Psychol.* 1985;4:449–488.
6. Perry CL, Klepp K-I, Sillers C. Community-wide strategies for cardiovascular health: the Minnesota Heart Health Program youth program. *Health Educ Res: Theory and Pract.* 1989;4:87–101.
7. Pentz MA, Dwyer JH, MacKinnon DP, et al. A multicommunity trial for primary prevention of adolescent drug abuse: effects on drug use prevalence. *JAMA.* 1989;261: 3259–3266.
8. Flay, BR. Mass media linkages with school-based programs for drug abuse prevention. *J Sch Health.* 1986;56:402–406.
9. Vartiainen E, Pallonen U, McAlister A, Puska P. Eight-year follow-up results of an adolescent smoking prevention program: the North Karalia Youth Project. *Am J Public Health.* 1990;80:78–79.
10. Blackburn H, Luepker RV, Kline FG, et al. The Minnesota Heart Health Program: a research and demonstration project in cardiovascular disease prevention. In: Matarazzo JD, Weiss SM, Herd JA, et al., eds. *Behavioral Health: A Handbook of Health Enhancement and Disease Prevention.* New York, NY: John Wiley; 1984:1171–1178.
11. Jacobs DR, Luepker RV, Mittelmark MB, et al. Community-wide prevention strategies: evaluation design of the Minnesota Heart Health Program. *J Chronic Dis.* 1986;39:775–788.
12. Cook TD, Campbell DT. *Quasi-experimentation: Design and Analysis Issues for Field Settings.* Palo Alto, Calif: Houghton Mifflin, 1979.
13. Perry CL, Jessor R. The concept of health promotion and the prevention of adolescent drug abuse. *Health Educ Q.* 1985;12: 169–184.
14. Perry CL, Murray DM. The prevention of adolescent drug abuse: implications from etioiogical, developmental, behavioral, and environmental models. *J Primary Prev.* 1985;6:31–52.
15. Klepp K-I, Halper A, Perry CL. The efficacy of peer leaders in drug abuse prevention. *J Sch Health.* 1986;56:407–411.
16. Perry CL, Griffin G, Murray DM. Assessing needs for youth health promotion. *Prev Med.* 1985;14:379–393.
17. Pechacek TF, Murray DM, Luepker RV, Mittelmark M, Johnson C, Shultz JM. Measurement of adolescent smoking behavior: rationale and methods. *J Behav Med.* 1984;7:123–140.
18. Komro KA, Kelder SH, Perry CL, Klepp KI. Effects of a saliva procedure on adolescent self-reported tobacco use and youth tobacco prevention outcomes. Manuscript under review.
19. Cummings SR, Richard RJ. Optimum cutoff points for biochemical validation of

smoking status. *Am J Public Health.* 1988; 78:5 574–575.
20. Murray DM, Hannan PJ. Planning for the appropriate analysis in school-based drug-use prevention studies. *J Consult Clin Psychol.* 1990;58:458–468.
21. Donner A, Birkett N, Buck C. Randomization by cluster: sample size requirements and analysis. *Am J Epidemiol.* 1990; 114:906–914.
22. Zucker DM. An analysis of variance pitfall: the fixed effects analysis in a nested design. *Educ Psychol Meas.* 1991;50:731–738.
23. *SAS/STAT Users Guide: Volume 2, GLM-VARCOMP, Version 6.* 4th ed. Cary, NC: SAS Institute Inc; 1989.
24. Flay BR, Koepke D, Thomson SJ, Santi S, Best JA, Brown KS. Six-year follow-up of the first Waterloo school smoking prevention trial. *Am J Public Health.* 1989;79: 1371–1375.
25. Murray DM, Pirie P, Luepker RV, Pallonen U. Five- and six-year follow-up results from four seventh-grade smoking prevention strategies. *J Behav Med.* 1989;12: 207–218.
26. Pirie PL, Murray DM, Luepker RV. Smoking prevalence in a cohort of adolescents, including absentees, dropouts, and transfers. *Am J Public Health.* 1988;78:176–178.
27. Hansen WB, Collins LM, Malotte CK, et al. Attrition in prevention research. *J Behav Med.* 1985;8:261–275.

Public Health Policy Forum Submissions Invited

The "Public Health Policy Forum," a new feature of the Journal, is intended to provide a presentation of divergent views on important public health policy issues that is more extensive than the Journal's format usually allows. Three kinds of material will appear in the Public Health Policy Forum: articles, not to exceed 4500 words; commentaries, not to exceed 2500 words; and editorials. Customarily, commentaries and editorials will be solicited by the Journal's Associate Editor or Editor. Articles for the Public Health Policy Forum will be selected from articles submitted to the Journal and, on occasion, from those submitted directly to the Associate Editor. In general, peer review will still be sought.

Those wishing to submit articles directly to the Associate Editor may do so at this address:

Bruce C. Vladeck, PhD
The United Hospital Fund
55 Fifth Avenue
New York, NY 10003-4392

Five copies of the article should be submitted simultaneously to the Journal's Washington, DC, office.

The scholarly merit and the scientific accuracy of all submissions will be considered. Additional criteria are relevance to an important policy issue, timeliness, and clarity and coherence of the policy argument. Brevity also helps considerably.

Locking Devices on Cigarette Vending Machines: Evaluation of a City Ordinance

Jean L. Forster, PhD, MPH, Mary E. Hourigan, MPH, and Steven Kelder, MPH

ABSTRACT

Objectives. Policymakers, researchers, and citizens are beginning to recognize the need to limit minors' access to tobacco by restricting the sale of cigarettes through vending machines. One policy alternative that has been proposed by the tobacco industry is a requirement that vending machines be fitted with electronic locking devices. This study evaluates such a policy as enacted in St. Paul, Minn.

Methods. A random sample of vending machine locations was selected for cigarette purchase attempts conducted before implementation and at 3 and 12 months postimplementation.

Results. The rate of noncompliance by merchants was 34% after 3 months and 30% after 1 year. The effect of the law was to reduce the ability of a minor to purchase cigarettes from locations originally selling cigarettes through vending machines from 86% at baseline to 36% at 3 months. The purchase rate at these locations rose to 48% at 1 year.

Conclusions. Our results suggest that cigarette vending machine locking devices may not be as effective as vending machine bans and require additional enforcement to ensure compliance with the law. (*Am J Public Health.* 1992;82:1217–1219)

Introduction

It has been well established that, in the United States, the onset of tobacco use occurs during adolescence.[1,2] School-based tobacco use prevention programs have shown limited long-term effects, leading to increased attention to strategies with potential to limit adolescents' access to cigarettes.[2] Reducing access to tobacco is an approach to tobacco use prevention that is favored by the general population as well.[3,4] Since 1989, at least 50 cities in Minnesota have adopted ordinances that restrict or prohibit cigarette vending machines because of the access to cigarettes they provide to teenagers.[3,5] Other cities and states are following Minnesota's lead in restricting youth access to tobacco through restrictions on vending machines and other policies.

In response, the tobacco and vending machine industries have proposed alternative policies that are less restrictive than vending machine bans, such as requiring cigarette vending machines to be retrofitted with locking devices that prevent cigarette purchase unless the lock is deactivated by an employee. A number of jurisdictions in the United States have adopted laws that require these devices on some or all cigarette vending machines.

Very little is known about the effectiveness of locking devices in limiting minors' ability to purchase cigarettes. Industry representatives maintain that such devices make vending machine purchase of cigarettes equivalent to over-the-counter sales by placing the sale under the control of an employee. Health advocates argue that (1) locking devices require extra enforcement to ensure that the devices are installed and operating and (2) the unlocking mechanism can be placed far away from the machine or even out of sight, eliminating the face-to-face encounter with the potential purchaser. Thus, a policy requiring locking devices is unlikely to be as effective as a vending machine ban in reducing the ability of minors to purchase cigarettes.

The only long-term experience with vending machine locking devices comes from Utah, which had a state law until 1989 requiring their installation. The experience in Utah was that the devices were rarely installed (1 of 22 machines was equipped with the device in one informal study), and, where installed, they were rarely operating (B. Neiger, written communication, February 1990). The Utah law was changed in 1989 to prohibit vending machine sales of cigarettes in most locations.

St. Paul, Minn, a city of approximately 270 000 people, adopted an ordinance requiring that locking devices be installed on all cigarette vending machines in the city effective July 1, 1990.[6] This policy provided the opportunity to measure the effectiveness of locking devices in a variety of business settings to limit minors' access to cigarettes. The purpose of this study was (1) to document the compliance of businesspeople with the new ordinance at 3 months and 1 year following the effective date and (2) to measure the effectiveness of such a policy in re-

Jean L. Forster and Steven Kelder are with the University of Minnesota. Mary E. Hourigan is with the Minnesota Department of Health, Minneapolis, Minn.

Requests for reprints should be sent to Jean L. Forster, PhD, MPH, University of Minnesota, School of Public Health, Division of Epidemiology, 1300 S Second St, Suite 300, Minneapolis, MN 55454.

This paper was submitted to the Journal October 4, 1991, and accepted with revisions March 10, 1992.

Editor's Note. See related editorial by Giovino et al. (p 1203) in this issue.

TABLE 1—Merchants' Compliance with Cigarette Vending Machine Locking Device Law

	Type of Business %			
	Restaurant (n = 24)	Bar or Liquor Store (n = 45)	Other (n = 8)	Total (n = 77)
Baseline				
Locking device	0	2	0	1
No locking device	100	98	100	99
3 months postlaw				
Locking device	50	42	13	42
No locking device	17	42	38	34
Over-the-counter sale only	0	11	25	9
No cigarette sales	33	4	25	16
1 year postlaw				
Locking device	42	56	13	47
No locking device	21	33	38	30
Over-the-counter sale only	4	11	25	10
No cigarette sales	33	0	25	13

ducing the ability of minors to purchase cigarettes from vending machines.

Methods

A list of all cigarette vending machine locations was obtained from the City of St. Paul in May 1990. Both the 1989 and 1990 lists were used since the licensing procedure changed in 1990 and some locations had not yet obtained their 1990 licenses. By comparing the two lists, 237 cigarette vending machine locations were identified. A random sample of 109 of these locations was selected using a random numbers table, with the goal of collecting data from 100 locations. Locations were subsequently excluded after visitation if there was no vending machine present, if the location was a private business with no access, or if the site was closed during baseline data collection. The total number of sites for which baseline data could be collected was 95. The vending machine locations included in this study were classified into restaurants (34%), bars and liquor stores (56%), and other (other stores, hotels; 10%). Cigarette purchase was attempted at these same locations at three points in time: just before the date the law was in effect, 3 months following baseline, and 1 year following baseline. For the results presented below, we included information for only those sites that were open for business on all three occasions when they were visited for the purchase attempt, so that we could document changes over time. This reduced the number of sites to 77.

All of the purchase attempts in this study were made by one individual, a girl who turned 15 years old just before the baseline observation. At each data collection point, this same girl was taken to each study site by an adult staff member. She dressed in age-appropriate clothing with no makeup. She attempted to buy cigarettes, and was debriefed immediately upon returning to the car by the driver. Items recorded in addition to whether she was able to purchase cigarettes included the date and time of day, the type of business, the type of sale (over the counter or vending machine), whether there was a locking device on the machine and whether it was operational, and whether anyone questioned her or commented about her age. All purchase attempts were completed between 3 PM and 7 PM.

Results

Merchants could comply with the law by installing a locking device, switching to over-the-counter sale of cigarettes, or discontinuing the sale of cigarettes altogether. Table 1 shows that 3 months after the law went into effect, over a third of the locations still had cigarette vending machines without a locking device. Businesses that mainly sold alcohol were least likely to comply, and restaurants were most likely. Merchants who did comply with the law were most likely to do so by installing locking devices, but a quarter of them either decided not to sell cigarettes at all or to switch to over-the-counter sales rather than install a locking device. By 1 year after the effective date, 30% of the merchants were still out of compliance with the law.

The second question regarding locking devices concerns the effectiveness of the law in limiting minors' ability to purchase cigarettes at the study locations. As shown in Table 2, the overall purchase success rate in these locations, all of which originally sold cigarettes through vending machines, decreased from 86% before the law went into effect to 30% at 3 months after the law was enacted; it subsequently rose to 48% at 1 year after the law. Businesses that installed locking devices on their vending machines were less likely to sell cigarettes to our teenage confederate than businesses with unlocked vending machines, but businesses that switched to over-the-counter sales were the least likely to sell cigarettes to minors, of those that still sold tobacco. Cigarette sales to our confederate declined the most at sites located in restaurants. The effect of the law deteriorated over the course of the year in all types of businesses, regardless of the merchants' choice of compliance, except in those that eliminated tobacco sales.

Comments about the age of the purchaser increased during the course of the study. At baseline, an employee questioned or commented about the confederate's age during 14% of the visits. After the locking device law had been in effect for 3 months, the rate of comments or questions rose to 55%, and then declined to 40% at 1 year.

Discussion

This study provides the first information about the effects of a policy that requires locking devices on cigarette vending machines. The results indicate that compliance with locking device laws is a problem. The implementation of the law was accompanied by a great deal of media attention. In addition, information about the requirements of the law was sent by the city licensing office to all vending machine license holders, and a city licensing official inspected each licensed location soon after the law went into effect, issuing a warning when the machine was not in compliance. Thereafter, each location was to be inspected annually, at the time of license renewal. Even 1 year after the law went into effect, 30% of the vending machines in this sample did not have locking devices. In contrast, two cities that have banned cigarette vending machines altogether show complete compliance, even almost 2 years after the ban took effect.[7] These data suggest that a locking device policy places an additional enforcement burden on city officials, thus reducing the potential effectiveness.

TABLE 2—Cigarette Purchase Success by Method of Compliance with Locking Device Law and Type of Business

	Baseline		3 Months Postlaw		1 Year Postlaw	
	%	No.	%	No.	%	No.
Type of sale						
Locking device	0	1	19	32	39	36
No locking device	87	76	65	26	91	23
Over-the-counter sale only	0	0	0	7	25	8
No cigarette sales	0	0	0	12	0	10
Type of business						
Restaurant	92	24	21	24	33	24
Bar or liquor store	80	45	33	45	56	45
Other	100	8	38	8	50	8
Total[a]	86	77	30	77	48	77

[a]Comparison of purchase success at 3 time points: $\chi^2 = 50.39$ (2 df); $P < .0005$; for Time 1–Time 2 difference, 95% CI = 41%, 71%; for Time 1–Time 3 difference, 95% CI = 24%, 52%.

Businesses that switched to over-the-counter sales of cigarettes or discontinued the sale of cigarettes added to the overall effectiveness of the law. The cost of retrofitting a vending machine to comply with the law is estimated at between $35 and $100, but only electronic vending machines can be modified.[8] These factors may have influenced some businesses to eliminate the machines. The overall effectiveness reflects low purchase success at sites that complied with the law (12% success at 3 months, 30% at 1 year) in combination with high success rates at sites that did not comply (65% success at 3 months, 91% at 1 year), providing evidence that low compliance reduces the potential effectiveness of a locking device policy.

Furthermore, effectiveness depends on employees checking age identification and refusing to deactivate the device if a minor attempts to purchase from the machine. This study shows that even when the locking device is installed, employees allowed purchase by our 15-year-old confederate 19% of the time at 3 months and 39% of the time at 1 year.

This study suggests that a locking device policy results in fewer locations selling cigarettes to minors than a policy of no restriction on cigarette vending machine sales but is probably less effective than the major policy alternative, a ban on cigarette vending machines. A follow-up of former vending machine locations in a community that instituted a ban shows only 8% of those locations selling tobacco to a 15-year-old confederate at 3 months and 15% at 15 months, compared with 30% and 48% at roughly the same time points in this study.[7] The community with the vending machine ban may have less tolerance for cigarette sales to minors, which could account for some of the difference in compliance. However, both cities are located in the same metropolitan area.

This study is limited by the fact that sites are drawn from a large urban city and may not represent the mix of locations or the responses of businesses in smaller cities or suburbs. Furthermore, the pre–post comparison design does not control for possible time trends, such as an increase in enforcement of the age of sale law. Such trends are likely to exaggerate the effectiveness of the locking device policy, and our results may represent the upper limit of the effectiveness of such a policy.

Controlling cigarette sales by vending machines is only one aspect of a coordinated local, state, and national policy effort proposed by many health professionals, policymakers, and researchers.[9-11] Our experience in Minnesota suggests that restricting vending machines is often the first step for communities, and the restrictions have both substantive and symbolic value.[7] Community leaders and businesspeople reason that enforcement of existing laws or adoption of new restrictions against over-the-counter sale of cigarettes is pointless as long as cigarettes are freely available in machines that do not allow discrimination by age of purchaser. Adoption of vending machine restrictions is often the first sign that a community is ready to treat cigarettes as a controlled and potentially dangerous product. It is important that these first steps toward controlling tobacco sales produce meaningful changes in availability, and do not simply add to the list of unenforced and unenforceable laws. □

Acknowledgments

This study was funded by a grant from the Park Nicollet Medical Foundation, Minneapolis, Minn.

We thank Lily Thompson, who was the 15-year-old buyer in this study; Scott Johnson for his help in implementing the study; and Harry Lando for helpful comments on the manuscript.

References

1. Johnston LD, O'Malley PM, Bachman JG. *Drug Use, Drinking, and Smoking: National Survey Results from High School, College, and Young Adult Populations.* Rockville, Md: National Institute on Drug Abuse; 1989.
2. *Reducing the Health Consequences of Smoking: 25 Years of Progress: A Report of the Surgeon General.* Washington, DC: US Dept of Health and Human Services; 1989.
3. Forster JL, Hourigan M, Weigum J. The movement to restrict children's access to tobacco in Minnesota. Presented to the Surgeon General's Interagency Committee on Smoking and Health; May 31, 1991; Washington, DC.
4. Jeffery RW, Forster JL, Schmid TL, McBride CM, Rooney BL, Pirie PL. Community attitudes toward public policies to control alcohol, tobacco, and high-fat food consumption. *Am J Prev Med.* 1990;6:12–19.
5. Forster JL, Hourigan M, McGovern P. Availability of cigarettes to underage youth in three communities. *Prev Med.* 1992; 21: 320–328.
6. City of St. Paul ordinances, Section 324.01, amended December 19, 1989.
7. Forster JL, Hourigan M, Pirie P, Weigum J. Feasibility of local policy changes to limit minors' access to tobacco. Unpublished manuscript.
8. *Preventing Tobacco Sales to Kids: What Can Communities Do?* St. Paul, Minn: Association for Non-Smokers Minnesota; 1990.
9. Forster JL, Klepp K-I, Jeffery RW. Sources of cigarettes for tenth graders in two Minnesota cities. *Health Educ Res.* 1989;4:45–50.
10. Altman DG, Foster V, Rasenick-Douss L, Tye JB. Reducing the illegal sale of cigarettes to minors. *JAMA.* 1989;261:80–83.
11. DiFranza JR, Norwood BD, Garner DW, Tye JB. Legislative efforts to protect children from tobacco. *JAMA.* 1987;257:3387–3389.
12. Centers for Disease Control. State laws restricting minors' access to tobacco. *MMWR.* 1990;39:349–353.
13. Koop CE. A parting shot at tobacco. *JAMA.* 1989;262:2894–2895.
14. National Commission on Drug-Free Schools. *Toward a Drug-free Generation: A Nation's Responsibility.* Washington, DC: US Dept of Education; 1990.

Adolescents' Multisubstance Use Patterns: The Role of Heavy Alcohol and Cigarette Use

Susan L. Bailey, PhD

ABSTRACT

Objectives. Knowledge about the roles that heavy alcohol and cigarette use play in patterns of concurrent substance use among adolescents is lacking despite studies showing that adolescent substance users are typically multisubstance users and that alcohol and cigarettes are commonly used heavily by those who use illicit substances.

Methods. The roles of increasing use and heavy first-time use of alcohol and cigarettes in multisubstance use patterns were examined in a cohort of 4192 secondary students who were surveyed three times over a 4-year period.

Results. When subsequent use patterns were compared for students who increased their levels of alcohol or cigarette use and those who initiated use for the first time but at heavy frequencies, analyses indicated that the former group was more likely to initiate the subsequent use of other substances and to maintain and increase use already initiated.

Conclusions. These results suggest that adolescents are likely to have been involved in a history of licit substance use characterized by increasing levels of use before progressing to and maintaining the use of other substances. Increasing frequencies of alcohol and cigarette use, therefore, may be markers for more serious patterns of substance use. (*Am J Public Health.* 1992;82:1220–1224)

Introduction

Studies focusing on the prevalence of adolescents' use of various types of substances have shown that substances are initiated in particular sequences.[1–7] A study conducted by Donovan and Jessor moved beyond focusing only on initiation to examine the position of heavy or problem alcohol use in the sequence of substance use among adolescents.[4] Results of their study suggest that heavy use of at least one licit substance may occur after the initiation of an illicit substance. The results of other studies suggest the importance of including heavy cigarette use when examining patterns of substance use among adolescents.[2,8–12]

Up to now, little attention has been given to the potentially different roles of heavy initiation or increasing use of licit substances or to the role of heavy cigarette consumption in general in adolescents' substance use careers. The main purpose of this analysis is to examine whether changes in levels of concurrent use of cigarettes and alcohol among youth can be used, over time, to predict subsequent substance use. In particular, this analysis evaluates whether heavy use of cigarettes and alcohol at an early point in time or increasing use of these substances over time presages more generalized substance use later on. Studies of substance use careers conclude that the longer the career, the more serious it is likely to be.[7,13] This conclusion suggests the hypothesis that increasing levels of use of licit substances during youth is more predictive of more serious patterns of use than is heavy initiation.

Methods

Subjects

The data used in this analysis were collected as part of two studies: the Study of Multilevel Substance Abuse Prevention for High-Risk Youth and the Study of Sequencing of Alcohol, Drugs, and Tobacco Among Youth. In these studies, a cohort of middle school (grades 6 to 8) students in a southeastern US county was followed through three rounds of data collection (T_1 – T_3) conducted from 1985 to 1990. Respondents were in grades 6, 7, and 8 (mean age = 12.3, SD = 0.98) at T_1; in grades 7, 8, and 9 (mean age = 13.7, SD = 1.0) at T_2, approximately 1 year after T_1; and in grades 9, 10, and 11 (mean age = 15.7, SD = 1.0) at T_3, approximately 2 years after T_2. Attrition was approximately 25% between each successive wave, and a total of 4192 students were included in all three rounds. Attrition analyses indicated that many of the substance users and heavier users included at T_1 subsequently were lost to the study. (See Table 1 for comparisons of mean levels of use at T_1 for those lost to and retained in the study.) Results reported here, therefore, are biased by the exclusion of a portion of the substance users. The results are also limited to students' self-reports of substance use. Analyses reported elsewhere, however, indicate that students were generally consistent in their reporting of use across rounds.[14] Additional details of the data are presented elsewhere.[15]

Susan L. Bailey is with the Center for Social Research and Policy Analysis at the Research Triangle Institute in North Carolina.

Requests for reprints should be sent to Susan L. Bailey, PhD, Center for Social Research and Policy Analysis, Research Triangle Institute, PO Box 12194, Research Triangle Park, NC 27709-2194.

This paper was submitted to the Journal July 17, 1991, and accepted with revisions May 5, 1992.

Editor's Note. See related editorial by Giovino et al. (p 1203) in this issue.

TABLE 1—Comparison of Students in a Cohort of Middle School Youth (Grades 6 to 8 at Round 1) Using *t* tests

Substances Used	Students Who Completed All Three Questionnaires		Students Who Completed Only the First Two Questionnaires			Students Who Completed Only the First Questionnaire		
	Mean	SD	Mean	SD	$P > \|t\|^a$	Mean	SD	$P > \|t\|^b$
Alcohol[c]	0.49	0.61	0.64	0.69	0.000	0.72	0.72	0.01
Cigarettes[c]	0.15	0.45	0.36	0.68	0.000	0.47	0.78	0.00
Inhalants[d]	0.06	0.28	0.11	0.32	0.002	0.13	0.33	0.27
Marijuana[b]	0.06	0.27	0.17	0.38	0.000	0.22	0.42	0.00
Other illicit drugs[d]	0.04	0.19	0.09	0.29	0.000	0.11	0.31	0.08

[a]Probability value for the *t*-test comparison between students who completed all three questionnaires and those who completed only the first two questionnaires.
[b]Probability value for the *t*-test comparison between students who completed only the first two questionnaires and those who completed only the first questionnaire.
[c]Categories for these variables: 0 = no use, 1 = light use, 2 = heavy use.
[d]Categories for these variables: 0 = no use, 1 = use.

Measures

The substance use measures focused on use in the year or, for one substance, use in the month prior to the survey. Students who reported a frequency of cigarette use between at least one pack a day and a half pack to one pack a week were coded as heavy users; light users were those who smoked less than a pack a month; and others were classified as abstainers during the period of interest. Students were coded as light alcohol users if they drank only once in the past year regardless of the number of drinks consumed, or if they drank any amount up to and including two drinks on the typical occasion (i.e., two to five times in the past year up to almost daily). They were coded as heavy drinkers if they drank three or more drinks on the typical occasion. Other substance use items included in the analyses measured past year use compared with no use of inhalants, marijuana, and other illicit drugs.

Results

Changes in Concurrent Use

Percentages of students reporting concurrent substance use patterns in the year prior to each round of data collection are presented in Table 2. For each round, patterns associated with percentages greater than 1 are reported. Percentages are reported separately for males and females and are compared to determine whether patterns differ by sex and whether subsequent analyses should be conducted separately for the two groups.

A comparison of common substance use patterns for males and females shows relatively little difference at all three rounds of data collection. The most consistent and sometimes largest difference concerns the heavy use of alcohol with no other substance use. At all three rounds, males were more likely than females to report this pattern. However, alcohol used heavily and in combination with other substances does not show a consistent difference by sex: sometimes females and sometimes males were more likely to report particular patterns involving heavy alcohol use. Other differences were small and not consistent across rounds. Because there were few differences in substance use patterns between males and females, subsequent analyses were conducted for both groups combined, and results are reported for the entire sample.

Percentages calculated for the total sample and reported in Table 2 show that the most commonly reported pattern at all three data collection points was no use of any substance; however, the proportion of abstainers declined by more than 20 percentage points from T_1 to T_3. At T_1, light alcohol use was by far the most common pattern of use and was a component of all but the two least common patterns among those that met the criteria to be reported (i.e., at least 1%). In the year prior to T_2, cigarettes emerged as a substance used either alone or in combination with other substances, typically at a heavy level when more than one other substance was used. When cigarettes were used heavily, they were most commonly used with alcohol, also at a heavy level and with all other substances. By T_3, heavy alcohol use appeared in the majority of the most common patterns. As for T_2, when alcohol and cigarettes were used heavily, they were commonly used in combination with other substances. Heavy alcohol and cigarette use, therefore, appeared to play a pivotal role in multisubstance use patterns. We now turn our attention to the roles of increasing use and heavy use initiation in these patterns.

Prevalences Relative to Transitions

Transitions were measured as changes in use (i.e., initiation or increased use) between T_1 and T_2 and were indicated by comparing reports of use at T_1 and T_2. Six such transitions were examined. These were (1) initiation of light alcohol use for those who were abstinent at T_1, (2) initiation of heavy alcohol use also for those who were abstinent at T_1, and (3) increase from light use at T_1 to heavy use at T_2, as well as (4) initiation of light cigarette use, (5) initiation of heavy cigarette use, and (6) increase from light to heavy cigarette use. The transitions were not mutually exclusive. Prevalences of other types of substance use subsequent to the transition (T_3) were examined for those who experienced each transition.

Percentages of students who initiated other substances during the 2-year period subsequent to the transitions of interest are presented in Figure 1. Generally, students who increased their frequency of alcohol use prior to T_2 were more likely than those experiencing other transitions involving alcohol to initiate subsequent use of other substances (with the exception of cigarettes). Too few cases of students who increased their frequency of cigarette use prior to T_2 were available to see if this pattern also holds for transitions in cigarette use.

The percentages of students who increased or decreased their frequencies of alcohol and cigarette use after the transitions of interest are presented in Figure 2. The graphs show that students were more likely to increase than to decrease their frequencies of use during this period.

TABLE 2—Percentage of Students in a Cohort of Middle School Youth (Grades 6 to 8 at T_1) Reporting Past Year Substance Use Patterns, by Students' Sex

Pattern	T_1 Male (n = 2014) %	T_1 Female (n = 2053) %	T_1 Total (n = 4067) %	T_2 Male (n = 1987) %	T_2 Female (n = 2039) %	T_2 Total (n = 4026) %	T_3 Male (n = 2000) %	T_3 Female (n = 2047) %	T_3 Total (n = 4047) %
No use of any substance	50.0	56.9	53.5	42.3	43.0	42.6	31.1	30.6	30.8
Light alcohol use only	27.7	24.2	25.9	21.3	22.7	22.0	17.6	21.1	19.4
Light cigarette use only	...[a]	...	...	1.7	2.3	2.0	0.9	1.3	1.1
Heavy alcohol use only	2.1	0.8	1.5	4.5	2.2	3.4	9.4	7.7	8.5
Inhalant use only	1.3	1.4	1.3	...	...	...	...	...	...
Light alcohol & light cigarette use	1.8	2.8	2.3	3.6	5.5	4.6	2.6	3.8	3.2
Light alcohol & heavy cigarette use	...	...	...	1.6	1.0	1.3	1.3	1.5	1.4
Light alcohol & inhalant use	2.6	2.4	2.5	1.2	1.4	1.3	...	...	...
Light alcohol & marijuana use	1.6	1.4	1.5	1.2	1.2	1.2	1.9	1.3	1.6
Heavy alcohol & light cigarette use	...	...	...	1.4	1.9	1.6	1.8	3.6	2.7
Heavy alcohol & heavy cigarette use	...	...	...	...	...	...	2.8	2.2	2.5
Heavy alcohol & marijuana use	...	...	...	...	...	...	4.0	1.7	2.8
Heavy alcohol, heavy cigarette, & marijuana use	...	...	...	2.4	1.2	1.8	3.9	3.8	3.8
Heavy alcohol, heavy cigarette, marijuana, & other illicit drug use	...	...	...	1.2	1.5	1.4	2.9	4.0	3.5
Heavy alcohol, heavy cigarette, inhalant, marijuana, & other illicit drug use	...	...	...	2.1	1.9	2.0	3.0	2.2	2.6
Other patterns	12.9	10.1	11.5	15.5	14.2	14.8	16.8	15.2	16.1
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Note. Percentages are presented separately for each round of data collection. T_1 refers to the 1985–1986 school year for urban schools and the 1986–1987 school year for rural schools. Respondents were in grades 6, 7, and 8 during this first round. T_2 refers to the 1986–1987 (urban) and 1987–1988 (rural) school years when students were in grades 7, 8, and 9. T_3 refers to the 1988–1989 and 1989–1990 school years when respondents were in grades 9, 10, and 11.

[a]Ellipses indicate less than 1% of the sample reported the pattern. The few students who did report the pattern were included in the "Other Patterns" category so columns would add to 100%. Otherwise, patterns are mutually exclusive.

Those who increased their frequency of cigarette use by T_2 were particularly unlikely to decrease their subsequent frequency of alcohol use.

Percentages of students who quit using substances subsequent to the transitions of interest at T_2 are presented in Figure 3. Generally, students were very likely to quit using substances during this period. The graphs presented in Figure 3 also show that students who increased their use of alcohol and cigarettes by T_2 were, in most cases, the least likely to quit the subsequent use of other substances. Those who initiated at light levels of use, on the other hand, were the most likely to quit other use later. This difference holds for subsequent discon-

tinuation of the same substance initiated or used more heavily by T_2. Students who initiated alcohol at light levels of use were most likely to quit subsequent alcohol use, followed by those who initiated at heavy levels, with those who increased use being the least likely to quit. Generally, the same pattern holds for cigarette use, although heavy cigarette initiators were slightly more likely than light initiators to quit later. Students who increased their cigarette use were the least likely of all students to quit cigarette use at a later time.

Discussion

As indicated in other studies,[1–7,16] results reported in this paper demonstrate that adolescents are more likely to initiate and use substances as they progress through middle school and into high school. Our results also suggest that the transition to heavy use of alcohol and cigarettes might be an important turning point in adolescents' substance use careers. When alcohol and cigarettes were used heavily by adolescents in our study, they were most often used in combination with illicit substances. These patterns were found for both males and females.

Our results also showed that it was less the level of alcohol and cigarette use that indicated subsequent multisubstance use than it was the progression to greater frequencies of use. Students who first initiated use at high levels were less likely to use these and other substances subsequently than were those who increased their frequencies of use. Students must have been involved in a longer history of use characterized by increasing levels of use in order to continue use and to progress to using other substances. These results support those reported by others, which conclude that the longer the substance use career, the more serious it is likely to be.[7,13] To these results, we add the conclusion that progression from lighter to heavier use of licit substances during adolescence likely plays a role in characterizing more serious substance use careers.

More research is needed to determine what predicts increasing frequencies of alcohol and cigarette use. No causal factors were tested here, and it may be found that the same factors that predict increasing use of licit substances also predict multisubstance use. Nevertheless, an adolescent's relative increase in the frequency of cigarette smoking or alcohol use appears to be more prognostic of multisubstance use

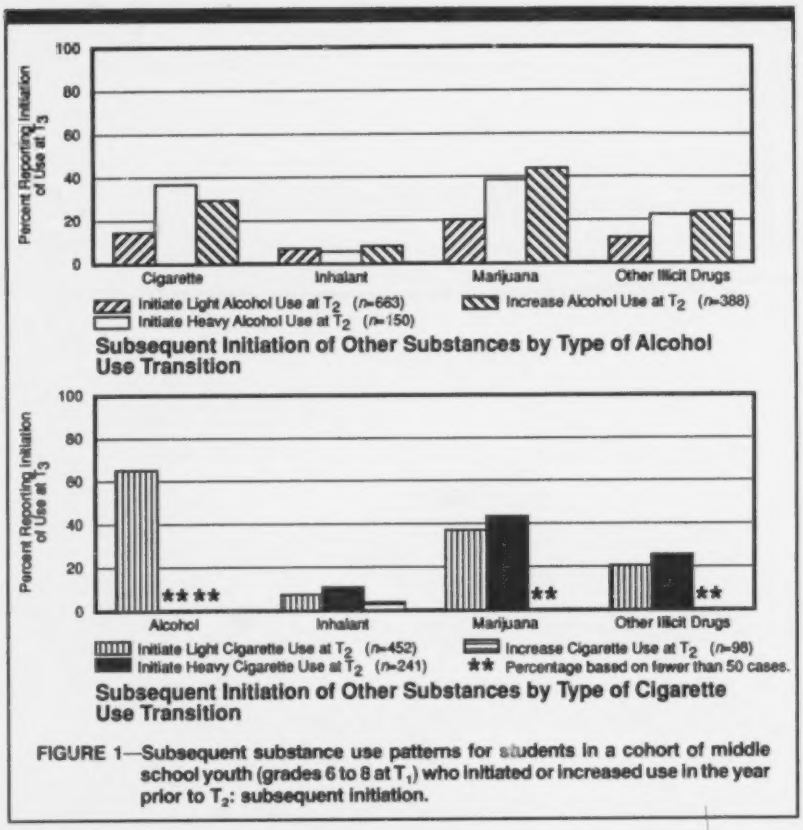


FIGURE 1—Subsequent substance use patterns for students in a cohort of middle school youth (grades 6 to 8 at T_1) who initiated or increased use in the year prior to T_2: subsequent initiation.

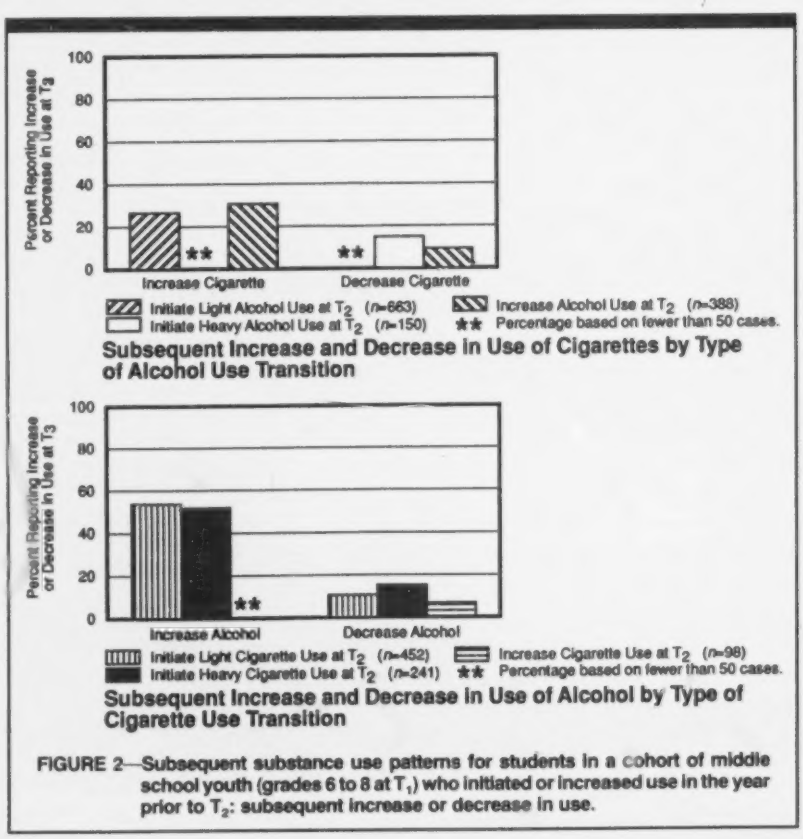

Percent Reporting Increase or Decrease in Use at T3
100
80
60
40
20
0
**
**
Increase Cigarette
Decrease Cigarette
Initiate Light Alcohol Use at T2 (n=663)
Initiate Heavy Alcohol Use at T2 (n=150)
Increase Alcohol Use at T2 (n=388)
** Percentage based on fewer than 50 cases.
Subsequent Increase and Decrease in Use of Cigarettes by Type of Alcohol Use Transition
Percent Reporting Increase or Decrease in Use at T3
**
Increase Alcohol
Decrease Alcohol
Initiate Light Cigarette Use at T2 (n=452)
Initiate Heavy Cigarette Use at T2 (n=241)
Increase Cigarette Use at T2 (n=98)
** Percentage based on fewer than 50 cases.
Subsequent Increase and Decrease in Use of Alcohol by Type of Cigarette Use Transition

FIGURE 2—Subsequent substance use patterns for students in a cohort of middle school youth (grades 6 to 8 at T_1) who initiated or increased use in the year prior to T_2: subsequent increase or decrease in use.

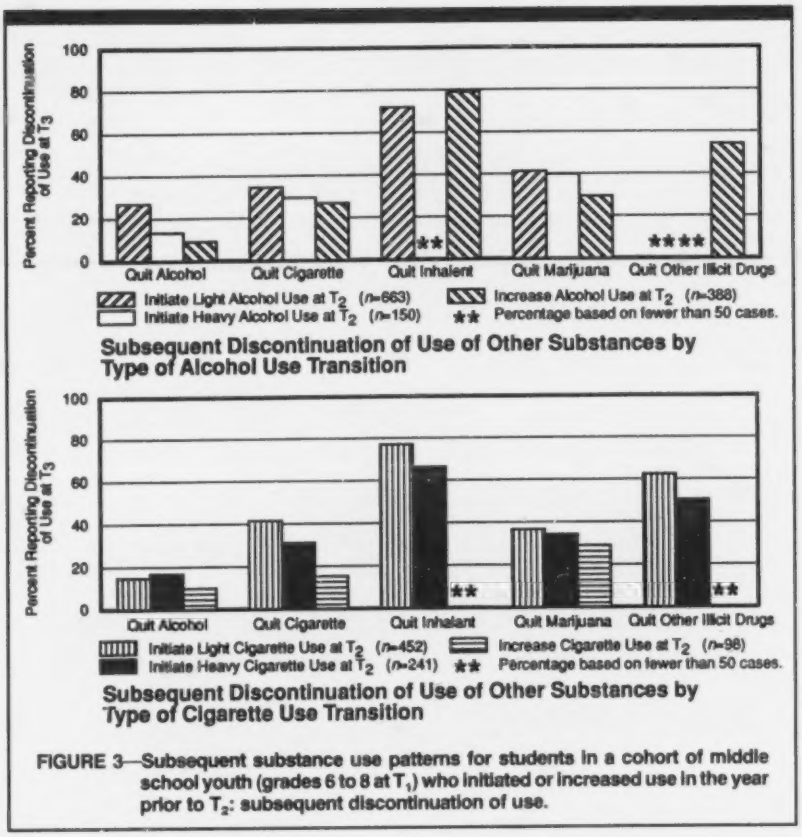


FIGURE 3—Subsequent substance use patterns for students in a cohort of middle school youth (grades 6 to 8 at T_1) who initiated or increased use in the year prior to T_2: subsequent discontinuation of use.

than is the mere frequency/quantity, even when frequency/quantity is high. Efficient targeting of early intervention efforts would focus on alcohol and cigarette users, and efforts would concentrate on preventing further progression to heavier levels of use and the often accompanying patterns of multisubstance use. □

Acknowledgments

The analysis reported in this paper was supported by the National Institute on Alcohol Abuse and Alcoholism Grant AA07666–03. The data collection was supported by Grant DA37971 from the National Institute on Drug Abuse. Dr. Robert L. Hubbard was the principal investigator for this earlier NIDA grant.

I wish to thank the late Teresa D. Crotts (1957–1991) and Catherine A. Boykin for their assistance in data analysis and manuscript preparation; and Dr. Christopher L. Ringwalt, J. Valley Rachal, Dr. Robert L. Flewelling, Richard S. Straw, and anonymous reviewers for their comments and editorial advice concerning successive drafts of this paper.

References

1. Kandel DB. Stages in adolescent involvement in drug use. *Science.* 1975;190:912–914.
2. Kandel D, Faust R. Sequence and stages in patterns of adolescent drug use. *Arch Gen Psychiatry.* 1975;32:923–932.
3. O'Donnell JA. Cigarette smoking as a precursor of illicit drug use. In: Krasnegn NA, ed: *Cigarette Smoking as a Dependence Process.* NIDA Research Monograph 23. Rockville, Md: National Institute on Drug Abuse; 1979;30–43. US Dept of Health and Human Services publication ADM 79–800.
4. Donovan JE, Jessor R. Problem drinking and the dimension of involvement with drugs: a Guttman scalogram analysis of adolescent drug use. *Am J Public Health.* 1983;73:543–552.
5. Yamaguchi K, Kandel DB. Patterns of drug use from adolescence to young adulthood: II. sequences of progression. *Am J Public Health.* 1984;74:668–672.
6. Raveis VH, Kandel DB. Changes in drug behavior from the middle to the late twenties: initiation, persistence, and cessation of use. *Am J Public Health.* 1987;77:607–611.
7. Fleming R, Leventhal H, Glynn K, Ershler J. The role of cigarettes in the initiation and progression of early substance use. *Addict Behav.* 1989;14:261–272.
8. Single E, Kandel D, Faust R. Patterns of multiple drug use in high school. *J Health Soc Behav.* 1974;15:344–357.
9. Block JR, Goodman N. Illicit drug use and consumption of alcohol, tobacco, and over-the-counter medicine among adolescents. *Int J Addict.* 1978;13:933–946.
10. Jessor R, Jessor SL. Theory testing in longitudinal research on marijuana use. In: Kandel DB, ed: *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* New York, NY: Hemisphere-Wiley; 1978:41–71.
11. Smith GM, Fogg CP. Psychological predictors of early use, late use, and nonuse of marihuana among teenage students. In: Kandel DB, ed: *Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues.* New York, NY: Hemisphere-Wiley; 1978;101–113.
12. Newcomb MD, Bentler PM. Frequency and sequence of drug use: a longitudinal study from early adolescence to young adulthood. *J Drug Educ.* 1986;16:101–120.
13. Kandel DB. Drinking and drug use among youth. *Annu Rev Sociol.* 1980;6:235–285.
14. Bailey SL, Flewelling RL, Rachal JV. The characterization of inconsistencies in self-reports of alcohol and marijuana use in a longitudinal study of adolescents. *J Stud Alcohol.* 1992;53. In press.
15. Bailey SL, Flewelling RL, Rachal JV. Predicting continued use of marijuana among adolescents: the relative influence of drug-specific and social context factors. *J Health Soc Behav.* 1992;33:51–66.
16. Bailey SL, Hubbard RL. Developmental changes in peer factors and the influence on marijuana initiation among secondary students. *J Youth Adolescence.* 1991;20:339–360.

Determinants of Passive Smoking in Children in Edinburgh, Scotland

Martin J. Jarvis, MPhil, David P. Strachan, MRCP, and Colin Feyerabend, PhD

ABSTRACT

Objectives. Using saliva cotinine as a quantitative marker, we examined the contribution of factors other than parental smoking to children's passive exposure to tobacco smoke.

Methods. Saliva specimens from a random sample of 734 7-year-old schoolchildren in Edinburgh, Scotland, were analyzed for cotinine. Their parents completed a questionnaire covering smoking habits and conditions in the home.

Results. A number of independent predictors of cotinine were identified in addition to the main one of smoking by household members. These predictors included home ownership, social class, day of the week, season of the year, number of parents present, crowding in the home, the number of children in the household, and sex. Cotinine was higher in children from less advantaged backgrounds, during winter, on Mondays, in girls, and when fewer other children were present. The effects were similar between children from nonsmoking and smoking homes.

Conclusions. Questionnaire measures of parental smoking are insufficient to fully characterize young children's exposure to passive smoking. Because socioeconomic variables contribute to measured exposure, passive-smoking studies that treat class as a confounder and control for it may be overcontrolling. (*Am J Public Health.* 1992;82:1225–1229)

Introduction

Studies examining the effects of passive smoking on children's health have mostly used questionnaire measures of parental smoking habits as their indicator of exposure to environmental tobacco smoke. This approach has given rise to an impressively consistent body of findings implicating passive smoking as a cause of respiratory illness in babies and young children.[1] However, commentators have pointed to the imprecision of questionnaire-based methods in quantifying exposure,[2,3] and in recent years an increasing number of studies have incorporated biochemical markers, principally cotinine, as objective measures. As a major metabolite of nicotine, cotinine is a sensitive indicator of the absorption of smoke products.[4] Several groups[5–11] have found dose-response relationships between the extent of smoking by household members and salivary or plasma cotinine concentrations in children, and associations between cotinine and low birth weight,[12] middle ear effusion,[13] and impairment of lung function[14] have been reported.

Despite the acknowledged importance of smoking by parents in determining their children's exposure to passive smoking, detectable concentrations of cotinine are found in the majority of children from homes where there are no smokers. In a study of 7-year-old children attending primary schools in Edinburgh, Scotland,[14] we noted that three quarters of the children from nonsmoking households had detectable salivary cotinine, and 10% of this group were in the upper two fifths of the distribution of measured tobacco-smoke exposure. This suggested that factors other than smoking by household members may be significant influences on exposure for some children. We have conducted further analyses of the data from the Edinburgh study in an attempt to identify determinants of exposure in addition to that of smoking by household members.

Subjects and Methods

A sample of 30 primary schools in Edinburgh (one third of all Edinburgh primary schools) was chosen at random and the parents of all children in the third year (aged 6 1/2 to 7 1/2 in September 1986) were contacted by postal questionnaire (n = 1095). The questionnaire inquired about respiratory symptoms in the child and conditions in the home. Usable replies were obtained from 1012 (92%), and written consent to further tests was obtained for 941 children (86% of the target sample). Twenty of these children left school before the survey began, and a further 20 were used in pilot studies. This left 901 children, of whom 892 participated in a clinical examination at school. Children were examined from January to June 1987 by Strachan. Measures taken included spirometry and a tympanogram. Full details of these procedures have been reported elsewhere.[13,14,15]

Martin J. Jarvis is with the Imperial Cancer Research Fund Health Behaviour Unit, London, England. David P. Strachan is with the Department of Public Health Sciences, St. George's Hospital Medical School, London. Colin Feyerabend is with the National Poisons Unit, New Cross Hospital, London.

Requests for reprints should be sent to Martin J. Jarvis, MPhil, Imperial Cancer Research Fund Health Behaviour Unit, Institute of Psychiatry, 4 Windsor Walk, London SE5 8AF UK.

This paper was submitted to the Journal August 7, 1991, and accepted with revisions March 17, 1992.

Editor's Note. See related editorial by Giovino et al. (p 1203) in this issue.

TABLE 1—Distribution of Saliva Cotinine (ng/ml) by Number of Smokers in Household

Percentiles	Number of Smokers in Household: None (n = 392)	One n = 223	Two or more (n = 119)	All subjects (n = 734)
10	ND	.40	1.30	ND
20	ND	.70	1.90	.10
25	ND	.80	2.20	.10
30	.10	1.00	2.70	.20
40	.10	1.46	3.30	.40
50	.20	1.80	4.40	.60
60	.30	2.74	5.10	1.20
70	.40	3.50	5.90	1.90
75	.40	3.90	6.50	2.80
80	.54	4.42	7.00	3.50
90	1.20	6.10	9.80	5.30
Mean (SD)	.46 (.84)	2.82 (2.82)	4.77 (3.11)	1.87 (2.85)
Geometric mean	.20	1.70	3.71	0.62

[a]None detected (< 0.1 ng/ml).

Children were asked to collect saliva in the mouth and to spit into a clean plastic container. A satisfactory specimen was obtained from 779 children (87% of those examined). The saliva samples were frozen within 8 hours of collection for later cotinine assay by gas-liquid chromatography, using a method with a detection limit of 0.1 ng/ml.[16]

On the questionnaire sent to parents, the question used to investigate smoking in the household was "How many people in your household smoke cigarettes?" Responses were coded as none, one, or two or more. Smoking by mothers, fathers, and others in the household was not separately ascertained. Social class was coded according to the British Registrar General's classification of occupations. Home ownership was categorized as rented (mainly low-income housing owned by the city) or owner-occupied. One-parent homes were defined as those where only one adult was reported as usually living in the household. An index of crowding was calculated by dividing the number of people in the household by the number of rooms (other than kitchen, bathroom, and toilets). Three levels were defined: low (less than one person per room); medium (one person); and high (more than one person). Day of the week when the examination was conducted was divided into Monday vs Tuesday through Friday. It was anticipated that this would permit identification of any differences in exposure between weekends and other times of the week. Season of the year of the examination was divided into winter (January to March) and spring (April to June).

Because the distribution of saliva cotinine values was highly skewed, a logarithmic transformation was applied to the values prior to statistical analysis, undetectable concentrations being assigned a value of 0.05 before transformation. We used multiple regression analyses to assess the independent association of predictor variables with measured saliva cotinine. Analyses of covariance were used to estimate the geometric mean cotinine concentration for each level of each predictor variable before and after adjusting for all other significant predictors.

Results

From a total of 779 children with cotinine data available, 6 had concentrations over 14.3 ng/ml, which has been suggested as a cut-off point to define active smoking[17] (values were 16.5, 17.1, 18.7, 21.0, 25.5, and 36.1 ng/ml). All of these children came from homes with at least one smoker, and it is possible that their high values were due entirely to breathing in other people's smoke. But since smoking is not unknown even in children as young as 7, and to avoid possible bias in assessing passive smoking effects, they were excluded from the analysis. A further 39 children had missing data on one or more variables of interest. This left a final sample of 734 children, of whom 16% had 2 or more smokers at home, 30% one smoker, and 53% no smokers. The equivalent percentages in the original sample of 1012 children for whom questionnaires were completed were 16, 31, and 53, indicating that the final sample was fully representative of the target population. Table 1, which gives percentiles of the distribution of saliva cotinine by household smoking, shows that cotinine concentrations were strongly influenced by the number of smokers living in the home. The median concentration was 0.2 ng/ml among children from nonsmoking homes, 1.80 ng/ml where one smoker was present, and 4.40 ng/ml where there were two or more smokers. The correlation between log cotinine and the number of smokers in the home was 0.71, indicating that this variable accounted for 51 percent of the variance in measured cotinine values.

Our approach to modeling saliva cotinine concentrations was to enter predictors with more than two levels as trend variables. In the case of household smoking, a slightly improved fit could be obtained by including separate dummy variables for one-smoker and two-smoker households. However, for simplicity of presentation, this variable was also treated as a linear trend. Table 2 gives the results of the multiple regression analysis. It shows that there were a number of other variables, in addition to household smoking, that reliably predicted cotinine concentrations. The indicators of socio-economic status—home ownership and social class—together accounted for a further 11% of the variance in cotinine. There was an inverse linear relationship between social class and cotinine, and, independently of this, children living in rented accommodations had higher concentrations than those from owner-occupied homes. Figure 1 shows that these effects were found among children from nonsmoking as well as smoking households. Children from one-parent households had concentrations considerably higher than other children's, and this effect was present whether or not the parent was a smoker. Both day-of-the-week and season-of-the-year effects were evident, concentrations being higher in winter than in spring, and in children examined on a Monday than on other days. Not unexpectedly, crowding in the home was associated with higher cotinine concentrations, but there was also a linear inverse relationship with the number of children living in the home. Finally, girls had higher measured exposure than boys.

Although statistically significant, the magnitude of these various effects was

generally small. For example, the difference in adjusted means between boys and girls was 0.14 ng/ml, similar to the day-of-the-week effect. However, there was a difference of nearly 0.5 ng/ml between children from professional homes and those with unskilled parents, and the single-parent effect was close to 1.0 ng/ml. When all predictors were entered, the percentage of variance in saliva cotinine accounted for was 68%. Variables that were examined but did not predict cotinine included the type of fuel used for cooking and heating, the presence of dampness in the home, and whether or not bedroom windows were reported as usually left open at night.

We examined the data for interaction effects between the main predictor variables, but none were identified. Because there were a number of predictors of cotinine besides the major one of the presence of smokers in the household, we analyzed how much of the limited variance in the group of children from nonsmoking homes (n = 392) these predictors could account for (Table 3). All predictors identified in the main analysis continued to show significant associations with cotinine. Overall they accounted for 33 percent of the observed variance. The direction and magnitude of these effects were generally similar to those found in smoking households (analysis not presented), confirming the lack of any significant interaction between these variables and parental smoking.

TABLE 2—Multiple Regression Analysis of Predictors of Saliva Cotinine Concentrations (n = 734)

R^2	Variable	No.	Unadjusted Geometric Mean	Adjusted Geometric Mean	β[a]	(95% CI)	t Value
.51	No. of smokers in household						
	0	392	0.20	0.26	3.57	(3.20, 3.98)	23.01
	1	223	1.70	1.28			
	2+	119	3.69	2.79			
.59	Home ownership						
	Owner-occupied	521	0.37	0.56	1.42	(1.16, 1.72)	3.48
	Rented	213	2.12	0.78			
.62	One-parent household						
	No	651	0.53	0.55	2.97	(2.32, 3.81)	8.58
	Yes	83	1.90	1.55			
.65	Social class[b]						
	I	90	0.18	0.41	1.19	(1.11, 1.27)	4.86
	II	199	0.27	0.49			
	III Non-manual	150	0.66	0.74			
	III Manual	144	1.17	0.67			
	IV and V	89	1.93	0.84			
.66	Month of examination						
	January–March	390	1.08	0.75	0.68	(0.51, 0.84)	−5.45
	April–June	344	0.33	0.49			
.67	No. of children in household						
	1	102	1.15	0.82	0.79	(0.72, 0.87)	−4.64
	2	392	0.57	0.66			
	3	172	0.61	0.53			
	4+	68	0.41	0.40			
.67	Persons per room						
	< 1	260	0.27	0.47	1.29	(1.16, 1.44)	4.54
	1	375	0.72	0.67			
	> 1	99	1.21	0.76			
.68	Sex						
	Boys	372	0.55	0.55	1.28	(1.11, 1.47)	3.34
	Girls	362	0.68	0.69			
.68	Day of examination						
	Monday	164	0.96	0.75	1.30	(1.09, 1.54)	2.89
	Tuesday–Friday	570	0.54	0.59			

Note. Variables listed in order of entry to a stepwise analysis. Regression coefficients and significance levels derived from a forced entry analysis in which all predictors were entered simultaneously. R^2 = .69; adjusted R^2 = .68; multiple r = .83.
[a]The regression coefficients shown have been exponentiated to represent the multiplicative effect on cotinine levels associated with a unit increase in the predictor variable.
[b]Coded according to the British Registrar General's classification of occupations.

Discussion

Our findings confirm the importance of smoking in the home in determining the extent to which young children breathe in other people's smoke, but they also show that a range of other factors also influence exposure. Smoking by adults living in the home accounted for 51 percent of the observed variance in saliva cotinine concentrations, while an additional 17 percent could be attributed to socioeconomic status, demographic factors, time at which the examination was conducted, and the extent of crowding in the home. The results—particularly for nonsmoking households—indicate that parental smoking habits as ascertained by questionnaire responses are inadequate to fully determine young children's exposure to environmental tobacco smoke.

Socioeconomic factors influenced the extent of exposure in two ways. First there was an association with the presence of smokers in the home (75% of children from blue collar, rented homes had one or more smokers at home compared with 29% of children from white collar, owner-occupied homes). But both social class and home ownership had further effects on measured exposure after controlling for the presence of smokers living in the home. Higher cotinine concentrations were associated with living in rented accommodations, and, in addition, concentrations showed an inverse linear relationship with social class as indexed by occupation of the head of household. This indicates that children from more deprived homes suffer additional exposure to passive smoking, possibly either through smoking by visitors to the home or by moving through a more smoke-polluted environment outside the home (stairwells, transport, shops, etc.).

Interpretation of the crowding effect observed would seem straightforward, implying simply that it is harder to avoid tobacco smoke the greater the number of people sharing the same living space. But the inverse association between number of children in the household and measured exposure was unexpected and harder to

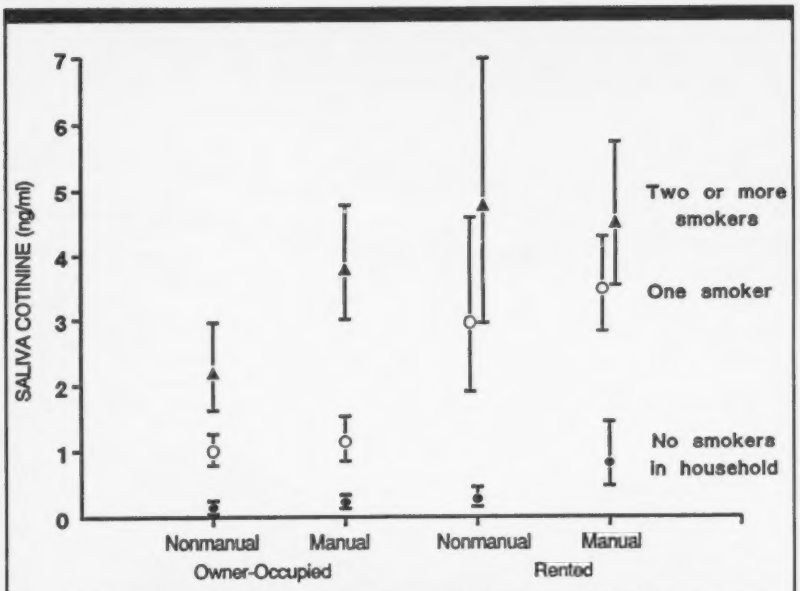


FIGURE 1—Geometric mean saliva cotinine concentrations (± 95% CI) by home ownership and social class among children from smoking and nonsmoking households.

TABLE 3—Multiple Regression Analysis Predicting Log Saliva Cotinine among Children with No Smokers in Household (n = 392)

	β[a]	95% CI	*t* Value
One-parent household	3.18	2.16, 4.68	5.49
Social class (trend I to IV/V)	1.22	1.10, 1.36	3.92
Season of examination	0.70	0.57, 0.87	−3.01
Home ownership	1.39	1.00, 1.93	2.01
Sex	1.32	1.07, 1.63	2.67
Persons per room	1.31	1.13, 1.53	3.59
Number of children in household	0.79	0.69, 0.91	−3.41
Day of examination	1.35	1.04, 1.75	2.26

Note. Multiple $r = .58$; adjusted multiple $R^2 = .33$.
[a]The regression coefficients shown have been exponentiated to represent the multiplicative effect on cotinine levels associated with a unit increase in the predictor variable.

account for. One explanation might be that each child functions as a kind of biological vacuum cleaner, so that the more children there are, the less smoke each can collect by breathing in. However, we found no similar effect for the number of adults in the household. More plausible is the notion that the more children there are, the more they will play with each other away from adults, and the more competition there will be for space close to parents.

Seasonal variation in saliva cotinine levels is likely to reflect better ventilation of the home in warmer, spring weather than in the winter. The Monday effect may reflect the fact that during the week the child spends a considerable portion of the time in the nonsmoking environment of the school, while during the weekend the home environment would predominate. Since cotinine has a half life in saliva of about 16 hours,[18] exposure during the weekend would have a considerable bearing on levels measured on a Monday. The existence of these effects of time of sampling suggests that studies will need to specify such factors carefully if their results are to be comparable.

The associations we have documented are clear-cut, but could they be due to factors other than passive smoking? Nicotine is not entirely specific to tobacco, being found for example in small amounts in peppers, aubergines, and potato skins.[19] Although it has been suggested that dietary factors may substantially influence cotinine levels,[20] a number of considerations strongly argue against this interpretation. It has been pointed out that people would have to eat huge amounts of these vegetables to have any measurable effect on their cotinine levels.[21] The putative dietary factors would also have to show exactly the same distribution in the population as parental smoking habits. Aside from this, the dietary hypothesis either offers no obvious explanation for observed associations (e.g., crowding, seasonal effects), or the direction of any postulated effect would be the opposite to that observed. Families in rented homes or with unskilled occupations would be expected to eat fewer vegetables, even though their cotinine concentrations are higher.

There was very considerable restriction on the range of cotinine values observed in children from nonsmoking homes. The median was 0.2 ng/ml and the 90th percentile 1.20 ng/ml. Despite this restriction, predictors emerged that were similar to those among smoking households, indicating that the reason measurable cotinine concentrations are found in children from nonsmoking homes is that in many cases they are still subject to a range of determinants of passive smoking. It would appear that the only practical way to determine their exposure is by means of a sensitive marker of actual smoke uptake.

Epidemiological investigations of passive smoking have treated variables such as social class and home ownership as potential confounders and have thus controlled for them in their statistical analysis of the relationship between exposure to passive smoking and disease endpoints. Our results indicate that in so doing they may be overcontrolling, because socioeconomic status and a number of other variables may each independently predict exposure to environmental tobacco smoke. Wherever possible, investigators should supplement or replace questionnaire data with quantitative measures of actual smoke uptake. □

References

1. *The Health Consequences of Involuntary Smoking: A Report of the Surgeon General.* Washington, DC: US Department of Health and Human Services, 1986.
2. Tager IB. "Passive smoking" and respiratory health in children—sophistry or cause for concern? *Am Rev Respir Dis.* 1986;133: 959–961.
3. Weiss ST. Passive smoking and lung cancer. *Am Rev Respir Dis.* 1986;133:1–3.
4. Jarvis MJ. Application of biochemical in-

take markers to passive smoking measurement and risk estimation. *Mutat Res.* 1989; 222:101–110.
5. Jarvis MJ, Russell MAH, Feyerabend C, Eiser JR, Morgan M, Gammage P, Gray EM. Passive exposure to tobacco smoke: salivary cotinine levels in a representative population sample of non-smoking schoolchildren. *Br Med J.* 1985;291:927–929.
6. Jarvis MJ, McNeill AD, Russell MAH, West RJ, Bryant A, Feyerabend C. Passive smoking in adolescents: one-year stability of exposure in the home. *Lancet.* 1987;1: 1324–1325.
7. Pattishall EN, Strope GL, Etzel RA, Helms RW, Haley NJ, Denny FW. Serum cotinine as a measure of tobacco smoke exposure in children. *Am J Dis Child.* 1985; 139:1101–1104.
8. Coultas DB, Howard CA, Peake GT, Skipper BJ, Samet JM. Salivary cotinine levels and involuntary tobacco smoke exposure in children and adults in New Mexico. *Am Rev Respir Dis.* 1987;136:305–309.
9. Greenberg RA, Haley NJ, Etzel RA, Loda FA. Measuring the exposure of infants to tobacco smoke. Nicotine and cotinine in urine and saliva. *N Engl J Med.* 1984;310: 1075–1078.
10. Greenberg RA, Bauman KE, Glover LH, et al. Ecology of passive smoking by young infants. *J Pediatr.* 1989;114:774–780.
11. Henderson FW, Reid HF, Morris R, et al. Home air nicotine levels and urinary cotinine excretion in preschool children. *Am Rev Respir Dis.* 1989;140:197–201.
12. Haddow JE, Knight GJ, Palomaki BS, McCarthy JE. Second-trimester serum cotinine levels in nonsmokers in relation to birth weight. *Am J Obstet Gynecol.* 1988; 159:481–484.
13. Strachan DP, Jarvis MJ, Feyerabend C. Passive smoking, salivary cotinine concentrations, and middle ear effusion in 7 year old children. *Br Med J.* 1989;298:1549–1552.
14. Strachan DP, Jarvis MJ, Feyerabend C. The relationship of salivary cotinine to respiratory symptoms, spirometry and exercise-induced bronchospasm in seven-year-old children. *Am Rev Respir Disease.* 1990; 142:147–151.
15. Strachan DP. Damp housing and childhood asthma: validation of reporting of symptoms. *Br Med J.* 1988;297:1223–1226.
16. Feyerabend C, Russell MAH. A rapid gas-liquid chromatographic method for the determination of cotinine and nicotine in biological fluids. *J Pharm Pharmacol.* 1990; 42:450–452.
17. McNeill AD, Jarvis MJ, West R, Russell MAH, Bryant A. Saliva cotinine as an indicator of cigarette smoking in adolescents. *Br J Addict.* 1987;82:1355–1360.
18. Jarvis MJ, Russell MAH, Benowitz NL, Feyerabend C. Elimination of cotinine from body fluids: implications for non-invasive measurement of tobacco smoke exposure. *Am J Public Health.* 1988;78:696–698.
19. Sheen SJ. Detection of nicotine in foods and plant materials. *J Food Sci.* 1988;53: 1572–1573.
20. Idle JR. Titrating exposure to tobacco smoke using cotinine—a minefield of misunderstandings. *J Clin Epidemiol.* 1990;43: 313–317.
21. Lee PN. Passive smoking and lung cancer association: a result of bias? *Hum Toxicology.* 1987;6:517–524.

Call for Abstracts for Epidemiology Late-Breaker Sessions

Exchange Session

The Epidemiology Section will sponsor a "late-breaker" Epidemiology exchange on Wednesday, November 11, 1992, during APHA's 1992 annual meeting in Washington, DC. The exchange will provide a forum for presentation of investigations or methods that have been conceived, conducted, and/or concluded so recently that members could not meet the deadline for abstract submission to other Epidemiology sessions. Abstracts should report on work conducted during the last 6–12 months.

Abstracts of fewer than 200 words (any format) and a stamped, self-addressed return envelope should be submitted to Polly A. Marchbanks, PhD, Chief, EIS Program, Epidemiology Program Office, Mailstop C08, Centers for Disease Control, Atlanta, GA 30333; (404) 639-3588.

Abstracts must be *received* by September 28, 1992. Decisions will be made by October 5, 1992.

Poster Session

The Epidemiology Section will again be sponsoring a late-breaker poster session at the annual meeting. The session will take place on Wednesday, November 11, 1992. Work completed in 1992 is eligible for consideration. Submit an abstract of fewer than 200 words (any format) and a *return envelope* to Cathey Falvo, MD, MPH, Graduate School of Health Science, Munger Pavilion, New York Medical College, Valhalla, NY 10595; tel. (914) 993-4323.

Abstracts must be *received* by October 2, 1992. Decisions will be mailed October 5, 1992. Students and recent graduates are particularly encouraged to submit abstracts.

Cigarette Smoking and Occupational Status: 1977 to 1990

Lirio S. Covey, PhD, Edith A. Zang, PhD, and Ernst L. Wynder, MD

ABSTRACT

Objectives. In this study we examined the relationship between occupational status and smoking habits in men and women during the period from 1977 to 1990.

Methods. Cigarette smoking and occupational history were obtained from 8045 men and women who served as controls for a hospital-based study of tobacco-related diseases.

Results. There was an association between increasing occupational status and tobacco exposure in men, but not in women. The quit rate increased over time in all sex-occupational groups except for male laborers, whose quit rate remained constant. Nicotine-dependent smokers are likely to find it difficult to quit. Male nicotine-dependent smokers were consistently found in greater numbers among blue collar workers throughout the study period. Initially, female nicotine-dependent smokers were more often found among blue collar workers, but in recent years became more frequent among white collar workers.

Conclusion. These trends provide clues to the future epidemiological distribution of lung cancer and other tobacco-related diseases. An understanding of gender differences in the occupational profile of cigarette smokers can provide guidelines for effective antismoking interventions. (*Am J Public Health.* 1992;82: 1230–1234)

Introduction

Ten years ago, we reported observations from a hospital-based epidemiological study that indicated a strong relationship between occupational status and smoking behavior in men.[1] Men in white-collar occupations, particularly those in professional-level jobs, were less likely to have ever smoked. Among individuals who did smoke, men in white-collar occupations were more likely to have quit, and among those who continued to smoke, white-collar workers were more likely to be smoking cigarette brands with lower tar and nicotine yields. The higher prevalence of cigarette smoking in male blue-collar workers than in male white-collar workers was also observed in representative samples of the US population.[2]

During the past decade, we have continued to obtain detailed smoking and occupational information from the same epidemiologic study that provided the database for our earlier study.[1] These data enabled us to perform an updated examination of the relationship between occupational status and smoking habits in both men and women, and to record the changes in smoking behavior by occupational category over the years 1976 to 1990. Additionally, using criteria suggested by the work of Fagerstrom,[3] we estimated the prevalence of "nicotine-dependent smokers" in each occupational group, by sex.

Methods

The data used in this analysis were collected as part of a large ongoing study of tobacco-related diseases that has been described elsewhere.[4] We included 8042 subjects (4985 male and 3057 female controls in the database) who were recruited between the years 1977 and 1990 from hospitals located in the New York, Long Island, and Chicago metropolitan areas in the United States and who did not have a history of tobacco-related diseases, such as cancers of the lung, oral cavity, larynx, esophagus, pancreas, kidney, or bladder; myocardial infarction, stroke, angina pectoris, or peripheral vascular disease; chronic bronchitis or emphysema. Among the admission diagnoses were malignancies (of the stomach, colon, prostate, and breast; leukemias; lymphomas; multiple myeloma; sarcomas), benign neoplasms, and acute or chronic conditions (e.g., cataracts, glaucoma, arthritis, fractures). Breast cancer in women (10%) and colorectal cancer in men (3%) were the most frequent diagnoses.

All data were collected by means of a detailed, standardized questionnaire filled out by trained interviewers; the questionnaire included information on demographic factors, smoking history, and occupation. Non-Whites and individuals younger than 40 years each constituted only approximately 5% of the population in the database, and they were excluded from the analyses.

Occupational Classification

Occupational classification was elicited by the question, "What has been your

Lirio S. Covey is with the Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY. Edith A. Zang and Ernst L. Wynder are with the Division of Epidemiology and Biostatistics, American Health Foundation, New York, NY.

Requests for reprints should be sent to Edith A. Zang, PhD, American Health Foundation, 320 East 43rd St, New York, NY 10017.

This paper was submitted to the Journal September 30, 1991, and accepted with revisions April 7, 1992.

Editor's Note. See related editorial by Giovino et al. (p 1203) in this issue.

INTENTIONAL
BLANK

TABLE 1—Characteristics of the Study Population

	1977–1980		1981–1984		1985–1990	
	Men	Women	Men	Women	Men	Women
	n = 1568	n = 1071	n = 1703	n = 956	n = 1714	n = 1030
Age, y						
40–49	16.2	15.5	14.9	15.6	16.3	13.6
50–59	36.4	37.2	34.0	34.7	31.2	29.5
60–69	36.0	35.3	40.0	35.2	39.4	39.4
70+	11.5	12.0	11.2	14.5	13.0	17.5
Education completed						
Elementary school	26.7	26.3	24.2	21.9	18.2	16.4
High school	25.8	37.4	30.1	41.3	27.2	38.5
College	47.6	36.3	45.7	36.7	54.6	45.1
Religion						
Protestant	21.4	20.6	21.8	19.9	21.5	18.9
Catholic	42.5	47.4	50.3	52.2	54.9	54.6
Jewish	29.3	28.3	22.0	24.2	21.2	24.7
Other	6.8	3.7	5.9	3.7	2.4	1.8
Marital status						
Single	5.2	7.2	7.6	6.0	6.6	5.4
Married	85.3	67.7	81.2	66.0	81.6	63.5
Divorced/separated	5.2	6.2	6.9	8.3	7.0	9.0
Widowed	4.3	18.9	4.2	19.8	4.8	22.0
City hospital						
Chicago	27.7	19.4	35.2	25.4	25.7	18.5
Long Island	19.4	12.5	17.7	16.6	20.0	20.9
New York	52.9	68.1	47.2	58.0	54.6	60.7
Occupation						
Professional/managerial	24.0	16.2	20.7	20.0	25.3	22.7
Manager	22.1	...	16.0	...	14.4	...
Sales/clerical	14.2	32.5	12.9	33.7	15.1	38.8
Craftsworker	18.4	...	27.0	...	28.6	...
Other blue-collar[a]	26.3	11.8	23.5	16.7	16.6	15.6
Homemaker	...	39.6	...	29.6	...	22.8

Note. All figures are percentages.
[a]Labor, operative, service.

usual occupation?" Responses were coded according to the categories employed by the US Bureau of the Census: professionals, managers, sales and clerical workers, craftsworkers, laborers, operatives, and service workers.[5] In addition, we included a separate code for homemakers, to designate women who had not generally been employed outside the home. Because of the small numbers, laborers, operatives, and service workers were grouped together as laborers in the analyses of data for men. For women, we combined professionals and managers into a single category, and craftsworkers, laborers, operatives, and service workers were classified as blue-collar workers.

Measures of Smoking Behavior

"Ever smokers" were defined as those who had ever smoked cigarettes at least once a day for 1 year. "Current smokers" included those who either were smoking at the time of the interview or had stopped less than 12 months earlier. Persons who smoked pipes or cigars, either alone or in addition to cigarettes, were excluded from the analyses. Rates of ever smoking and current smoking were both calculated using the total sample size in each sex-specific occupational group as the denominator.

"Quitters" were defined as previous smokers who had not smoked during the previous year, and the proportion of quitters was calculated using the total number of ever smokers in each sex-specific occupational group as the denominator. Additionally, we calculated the proportions of smokers who smoked more than 30 cigarettes per day, of those who smoked cigarettes yielding more than 20 milligrams of tar, and of nicotine-dependent smokers, each time using the total number of current smokers in each sex-specific occupational category as the denominator.

Drawing from the literature on the Fagerstrom Tolerance Questionnaire,[2] we defined nicotine-dependent smokers as current smokers who smoked more than 25 cigarettes per day and who lit up in the morning within 30 minutes of awakening. It has been shown that these two behaviors are both highly associated with biochemical measures of smoking—for example, cotinine and nicotine in the serum[6]—and are most characteristic of smokers who are unable to quit.[7]

Data Analysis

The percentages of individuals in each of the above smoking categories were calculated by occupation and sex.

TABLE 2—Specific Smoking Behaviors in Men, by Occupational Status, in Three Time Periods

	Time 1 (1977–1980) n = 1568	Time 2 (1981–1984) n = 1703	Time 3 (1985–1990) n = 1714	Difference Time 3 – Time 1
Ever smokers				
Professional	57.2	62.3	59.6	+2.4
Managerial	63.9	72.8	68.8	+4.9
Sales/clerical	62.6	71.7	74.1	+11.5
Craftsworker	72.3	72.3	71.6	−0.7
Laborer/operative	70.0	80.8	81.4	+11.4
Current smokers				
Professional	21.5	21.3	16.4	−5.1
Managerial	27.2	29.8	24.7	−2.5
Sales/clerical	28.8	28.8	27.8	−1.0
Craftsworker	37.4	33.1	26.1	−11.3
Laborer/operative	33.1	38.8	39.0	+5.9
Quitters[a]				
Professional	62.3	65.9	72.5	+10.2
Managerial	57.5	59.0	64.1	+6.6
Sales/clerical	54.0	59.9	62.5	+8.5
Craftsworker	48.3	54.2	63.5	+15.2
Laborer/operative	52.4	52.0	52.5	+0.1
Smoking ≥30 cigarettes/d[b]				
Professional	45.1	41.6	40.9	−4.2
Managerial	48.4	48.0	43.5	−4.9
Sales/clerical	52.5	51.0	54.7	+2.2
Craftsworker	51.2	49.7	46.4	−4.8
Laborer/operative	53.2	48.6	51.3	−1.9
Smoking ≥20 mg tar/cigarette[b]				
Professional	14.8	10.7	7.0	−7.8
Managerial	16.0	9.9	11.5	−4.5
Sales/clerical	12.5	9.5	13.9	+1.4
Craftsworker	24.1	20.4	14.8	−9.3
Laborer/operative	15.3	23.9	15.3	0.0
Nicotine-dependent[b]				
Professional	33.3	26.7	35.2	+1.9
Managerial	35.1	43.2	31.2	−3.9
Sales/clerical	42.2	39.7	47.2	+5.0
Craftsworker	41.7	47.4	45.3	+3.6
Laborer/operative	46.0	40.0	51.4	+5.4

Note. All figures are percentages.
[a]Based on ever smokers.
[b]Based on current smokers.

We examined temporal changes in smoking by dividing the data into three time intervals: 1977 to 1980, 1981 to 1984, and 1985 to 1990. Although the time intervals correspond with progressively expanded and modified versions of the survey instrument, none of the revisions involved the variables analyzed in this study.

The three geographic areas (Chicago, Long Island, and New York) that served as the sources for recruitment of our sample did not yield substantially different distributions by age, occupation, and smoking exposure. We therefore combined them in subsequent analyses.

Results

Demographic Characteristics

The characteristics of the study population are summarized in Table 1. In men, the age distribution was relatively consistent over time; more than 80% of the sample were older than 50 years of age, and roughly half were older than 60 years of age. The proportion of college-educated men in relation to those with only an elementary education increased slightly over time. The sample included more Catholics than men of other religions, and more than 80% of the men were married. By geographic location, New York hospitals accounted for about 50% of the sample. About half of the men held white-collar jobs.

The demographic distributions of the women in the study resembled those of the men for all three time periods, except that fewer women were college educated and women were more likely than men to be widowed. Also, there were fewer women than men in professional and managerial occupations, but more in the sales/clerical category. In line with nationwide trends, the proportion of homemakers decreased over time.

The Relationship Between Smoking Behavior and Occupation

Men (Table 2). In each of the time periods, professional men were less likely to have ever smoked than were those in other occupational groups, and all groups except craftsworkers showed an increase in ever smokers from time 1 to time 3. The largest change, which occurred between time 1 and time 2, was seen among sales/clerical workers and laborers.

Professionals consistently had the highest quit rates, and men in blue-collar occupations had the lowest quit rates. Over time, quit rates increased and the proportion of current smokers decreased in all occupational categories with the exception of laborers. Blue-collar workers were twice as likely as professionals to be current smokers.

A less consistent association was found between heavy smoking (≥30 cigarettes per day) and occupation. Professionals were least likely to be heavy smokers, but sales/clerical workers surpassed blue-collar workers in heavy smoking at times 2 and 3, and they were the only group that showed an increase in heavy smoking over time.

In all categories except sales/clerical workers and laborers, the proportion of smokers who smoked high-tar-yield cigarettes decreased over time. The proportion of such smokers was consistently lower among white-collar workers.

Nicotine dependence was least common among professionals, particularly in time 2, and most common among blue-collar workers. In contrast to the general decline in smoking exposure (except among laborers), the proportion of dependent smokers showed an overall increase over time.

A noteworthy finding is the reduction in smoke exposure among craftsworkers, the most highly skilled blue-collar group,

among whom the proportion of current smokers declined while the quit rate continued to increase.

Women (Table 3). Fewer women than men smoked, and their smoking behavior was less clearly associated with occupation. Homemakers did not consistently resemble members of any other occupational category in their smoking patterns.

In contrast to men, proportionately fewer women white-collar workers had ever smoked. Blue-collar workers and homemakers showed a notable increase in this category over time.

Women's patterns of quitting were similar to those of men: quit rates were generally highest among professionals and managers and lowest among women in blue-collar jobs. Although homemakers resembled blue-collar workers in their increasing frequency of ever smoking, the proportion of quitters among homemakers was similar to that among white-collar workers. Overall, quit rates increased steadily over time among women. In accord with the trends in quit rates, professional women and homemakers were less likely to be current smokers than were women in blue-collar jobs. As with men, the numbers of current smokers diminished over time.

No consistent trend in the frequency of heavy smokers by occupation was evident among women. There was a general decline in heavy smoking, but this decline did not occur among those in sales and clerical positions.

Fewer women than men smoked high-tar-yield cigarettes, and the proportion of such smokers declined over time in all occupations. Women in blue-collar jobs were more likely to smoke the high-tar-yield brands, but they also showed the largest decrease in this behavior over time.

There was no clear relationship between occupation and nicotine dependence among the women in the sample, except that the proportions of dependent smokers increased in the white-collar occupations but decreased among blue-collar workers and homemakers.

TABLE 3—Specific Smoking Behaviors in Women, by Occupational Status, in Three Time Periods

	Time 1 (1977–1980) n = 1071	Time 2 (1981–1984) n = 956	Time 3 (1985–1990) n = 1030	Difference Time 3 – Time 1
Ever smokers				
Professional/managerial	50.3	51.3	53.4	+3.1
Sales/clerical	50.3	50.6	50.8	+0.5
Blue-collar[a]	32.5	48.8	42.9	+10.4
Homemaker	41.0	51.2	48.9	+7.9
Current smokers				
Professional/managerial	23.7	28.3	16.2	−7.5
Sales/clerical	27.9	25.5	22.5	−5.4
Blue-collar[a]	25.4	28.1	22.4	−3.0
Homemaker	23.4	23.7	18.3	−5.1
Quitters[b]				
Professional/managerial	52.9	44.9	69.6	+16.7
Sales/clerical	44.6	49.7	55.7	+11.1
Blue-collar[a]	22.0	42.3	47.8	+25.8
Homemaker	43.1	53.8	62.6	+19.5
Smoking ≥30 cigarettes/d[c]				
Professional/managerial	25.3	30.6	24.0	−1.3
Sales/clerical	26.3	30.7	32.0	+5.7
Blue-collar[a]	34.2	26.9	27.5	−6.7
Homemaker	28.7	26.4	23.5	−5.2
Smoking ≥20 mg tar/cigarette[c]				
Professional/managerial	9.8	7.4	2.6	−7.2
Sales/clerical	8.3	8.5	4.4	−3.9
Blue-collar[a]	21.9	15.6	8.3	−13.6
Homemaker	11.1	6.0	2.3	−8.8
Nicotine-dependent[c]				
Professional/managerial	17.1	22.2	26.3	+9.2
Sales/clerical	17.5	35.4	35.6	+18.1
Blue-collar[a]	31.3	22.2	16.7	−14.6
Homemaker	26.3	26.9	20.9	−5.4

Note. All figures are percentages.
[a]Craftsworkers, laborers, operatives combined.
[b]Based on ever smokers.
[c]Based on current smokers.

Discussion

Although our study population consisted of hospitalized patients, the information obtained on their smoking behavior and usual occupation referred to their activities prior to their diagnoses. It is unlikely that these variables would have been influenced by their illness. An additional possibility of bias would be due to key differences in the populations sampled at the three time periods. However, as shown in Table 1, the demographic characteristics were approximately the same except for the overall increase in the educational level of the latest sample, which reflects a nationwide trend.

The results can be summarized as follows:

1. The apparent strong association between occupation and cigarette smoking among men that we reported 10 years ago continues.

2. Smoking behavior is much less related to occupation among women than among men.

3. The proportion of smokers who are most likely to have difficulty when attempting to quit, that is, nicotine-dependent smokers, is lowest among men in professional and managerial occupations. In women, dependent smokers were more frequently found among blue-collar workers and homemakers during the first period of observation (1977 to 1980). Over time, however, the proportion of nicotine-dependent smokers decreased among blue-collar workers and increased among women in white-collar occupations.

Among men, higher status occupations were generally associated with lesser smoking exposure as estimated by quit rate, current smoking prevalence, average tar yield per cigarette, and nicotine dependence. In contrast, although women in higher level occupations were more likely to have ever smoked, they were also more likely to have quit and to smoke lower-tar-yield brands than were those em-

ployed in lower level occupations. Homemakers, undoubtedly a heterogeneous group with regard to sociodemographic status, showed an irregular pattern in their smoking behavior as well as in changes over time; in some ways they resembled professional women and in others they resembled blue-collar workers.

Our results also indicate that women included fewer ever smokers than men, but that women had somewhat lower quit rates. Consequently, the proportions of current smokers were not much lower among women than among men, although women smoked fewer cigarettes per day, were less likely to smoke high-tar-yield brands, and were less likely to be dependent smokers. In a related paper, it has been shown that the women in this sample began to smoke later in life than did men and that this delayed onset of smoking was subsequently associated with a lower number of cigarettes smoked per day.[8] These findings suggest that although the incidence of tobacco-related diseases may continue to rise among women, as long as present smoking patterns prevail it will not reach the peak levels recorded for men.

The higher prevalence of tobacco use among men in blue-collar jobs is significant in light of the fact that these occupational groups are also more likely to be exposed to other hazardous substances through their work. Although broad estimates of lung cancer risk attributable to occupations range from 4% to 36%, with a mean of 13%, the estimated proportion of risk attributable to cigarette smoking is far greater, about 85%.[9] Consequently, increased efforts at reducing smoking exposure would contribute substantially to the overall reduction in the risk of lung cancer and other diseases related to both smoking and occupational exposures.

With the exception of male laborers, the proportion of current smokers in all occupations and in both sexes decreased over time, accompanied by a concurrent increase among current smokers in the frequency of nicotine dependence. This increasing proportion of individuals who are unable to quit points up the continuing need for efficacious smoking cessation intervention programs, particularly those targeted to reach the hard-core, nicotine-dependent smoker.

The difference by gender in the occupational profile of dependent smokers, who are predominantly blue-collar workers among men but predominantly white-collar workers among women, is intriguing. An understanding of the apparent interaction between gender and occupational status with regard to dependent smoking may lead to more efficacious smoking prevention and smoking cessation programs in the workplace. □

Acknowledgment

This investigation was supported by grant CA32617 from the National Cancer Institute, US Department of Health and Human Services.

References

1. Covey LS, Wynder EL. Smoking habits and occupational status. *J Occup Med.* 1981;23: 537–542.
2. *Reducing the Health Consequences of Smoking: 25 Years of Progress.* A report of the Surgeon General. Rockville, Md: Centers for Disease Control, Office on Smoking and Health; 1989. US Dept of Health and Human Services publication CDC 89-8411.
3. Fagerstrom K, Schneider NG. Measuring nicotine dependence: a review of the Fagerstrom Tolerance Questionnaire. *J Behav Med.* 1989;12:159–182.
4. Wynder EL, Stellman SD. Comparative epidemiology of tobacco-related cancers. *Cancer Res.* 1977;37:4608–4622.
5. US Bureau of the Census. Table 227. Detailed characteristics. *Final Report PC(11)-D1.* United States Summary. Washington, DC: US Govt Printing Office; 1973.
6. Heatherton TF, Kozlowski LT, Flecker RC, et al. *Br J Addiction.* In press.
7. Kabat GC, Wynder EL. Determinants of quitting smoking. *Am J Public Health.* 1987; 77:1301–1305.
8. Taioli E, Wynder EL. The importance of age of starting smoking. *N Engl J Med.* In press.
9. Schoenberg JB, Stemhagen A, Mason TJ, et al. Occupation and lung cancer risk among New Jersey white males. *JNCI.* 1987;79:13–21.

Predictors of Smoking Relapse after Noncardiac Surgery

ABSTRACT

Objectives. We studied 235 smokers who were admitted for major noncardiac surgery and who survived hospitalization to determine the frequency and predictors of smoking relapse 1 year after hospitalization.

Methods. We analyzed data collected prospectively from the Study of Perioperative Ischemia. We determined smoking status by self-report.

Results. At 12 months posthospitalization, 142 (76%) of the 186 surviving patients had resumed smoking. Significant independent predictors of relapse at 12 months were a history of alcohol abuse and a history of heavy cigarette smoking ($P < 0.01$ and $P < 0.05$, respectively).

Conclusions. Since smokers commonly undergo surgery, the perioperative period provides a valuable window of opportunity for smoking cessation interventions designed to decrease the high rate of smoking relapse after surgery. (*Am J Public Health.* 1992;82:1235–1237)

Joel A. Simon, MD, MPH, Warren S. Browner, MD, MPH, and Dennis T. Mangano, PhD, MD, for the Study of Perioperative Ischemia (SPI) Research Group

Introduction

About one in three adult men and one in four adult women in the United States continue to smoke despite the current ongoing public health efforts to encourage cessation.[1] As many as 90% of regular smokers would like to quit.[2] Because smoking represents the single most important preventable cause of loss of life in the United States, designing and implementing successful smoking cessation interventions is of obvious public health importance.[3] More than 25 million surgeries are performed each year in the United States,[4] providing a captive audience of smokers who usually must quit smoking, at least temporarily during the perioperative period.

The Study of Perioperative Ischemia (SPI) enrolled men with coronary disease or at high risk for it, who were undergoing elective noncardiac surgery.[5] Approximately half of these men were smokers. Since all smoking subjects were abstinent during the postoperative period, all were de facto quitters. To determine the epidemiology of smoking relapse among these men, our study followed these men prospectively and ascertained their quit rates at 12 months.

Methods

Subjects

The full details of the SPI have been described elsewhere.[5] The cohort consisted of 474 men at the San Francisco Veterans Affairs Medical Center who were scheduled for elective noncardiac surgery requiring general anesthesia and were enrolled consecutively between January 1987 and September 1989. Of these, 241 (51%) were self-reported smokers at the time of enrollment. Excluded from the analysis were 6 smokers who died during the initial hospitalization, 41 who died during the subsequent year, and 8 who were lost to follow-up. This left 186 subjects available for analysis at 12 months.

Measurements

Age, race, and marital status were ascertained after admission. Any history of coronary artery disease (CAD), congestive heart failure, valvular heart disease, peripheral vascular disease, cerebrovascular disease, treated hypertension, diabetes mellitus, or recent surgery was recorded. Patients were questioned regarding the use of bronchodilators, which this study used as a surrogate for the diagnosis of obstructive airway disease.

The current number of packs smoked per day and the number of years of smoking at this level were recorded on admission. Pack-year history of smoking was then estimated. Patients were also asked if they had a history of alcohol abuse, but there was no attempt to quantify level of alcohol use.

Smokers undergoing major vascular and intrathoracic surgery were considered

Joel A. Simon and Warren S. Browner are with the General Internal Medicine Section and Dennis T. Mangano is with the Department of Anesthesia of the Veterans Affairs Medical Center in San Francisco. The authors are also with the University of California in San Francisco.

Requests for reprints should be sent to Joel A. Simon, MD, MPH, General Internal Medicine Section (111A1), Veterans Affairs Medical Center, 4150 Clement St, San Francisco, CA 94121.

This paper was submitted to the Journal December 16, 1991, and accepted with revisions June 16, 1992.

Editor's Note. See related editorial by Giovino et al. (p 1203) in this issue.

TABLE 1—Characteristics of the Study Subjects

	Subjects (n = 186)	
	No.	%
White	143	77
Black	28	15
Married	33	18
Coronary artery disease	80	43
Congestive heart failure	20	11
Peripheral vascular disease	93	50
Cerebrovascular disease	40	22
Diabetes mellitus	29	16
Alcohol abuse	102	55
Hypertension	91	49
Bronchodilator use	14	8
Type of surgery: Major vascular/intrathoracic	86	46
In-hospital cardiac event	12	7

Note. Subject characteristics calculated by mean ± SD included: age (y), 65 ± 8.7; current tobacco use (packs per day), 1.1 ± .66; pack-year history (n = 185), 40.8 ± 12.1; body mass index (kg/m^2; n = 177), 24.7 ± 4.2; and postoperative days, 15.4 ± 15.3.

together and compared with smokers undergoing all other operative procedures. Because major vascular and intrathoracic surgeries are often performed for diseases directly related to tobacco use, we believed that these smokers might be more likely to remain quitters postoperatively than those undergoing other surgical procedures. The number of postoperative days in the hospital was used to estimate length of hospital stay and initial tobacco abstinence.

Twelve-month postoperative follow-up of patients was performed by direct interview and included self-reported smoking status.

Statistical Analysis

Logistic regression analyses were performed with SAS Proc Logist.[6] Variables that did not meet the .10 significance level in the univariate models were not entered into the multivariate model. Odds ratios (OR) with 95% confidence intervals (CI) were calculated. Two-tailed t tests and Mann-Whitney U tests were done to compare continuous data, and χ^2 tests were performed for the analysis of categorical data.

Results

Our cohort was predominantly unmarried and White (Table 1). Alcohol abuse and peripheral vascular disease were common. Half of our patients were taking antihypertensive medications. The study subjects were, on average, moderate smokers at the time of entry, with 40 pack-years of tobacco use.

At 12 months posthospitalization, the rate of relapse was 76%. Smokers with a history of alcohol abuse were more likely than non-alcohol abusers to resume smoking (84% vs 67%, $P < .01$). Other significant univariate predictors of relapse at 12 months included CAD, pack-year history of smoking, and duration of hospitalization. Among CAD patients, 70% relapsed as compared with 81% among non-CAD patients ($P < .10$). Smokers with hospital stays of 2 or more weeks were less likely to relapse than smokers with shorter stays (66% vs 81%, $P < .05$), and relapsers had on average 5 fewer days of hospitalization ($P < .05$). Smokers with 50 or more pack-years of smoking had a relapse rate of 87% compared with 73% for the 30- to 49-pack-year smokers and 64% for smokers with under 30 pack-years ($P < .05$).

Multivariate analysis disclosed that only pack-year history of smoking and history of alcohol abuse were statistically significant independent predictors of relapse at 12 months (Table 2). Using the logistic coefficients to estimate the risk of relapse, a 40 pack-year smoker with a history of alcohol abuse had a 94% relapse probability compared with a 71% probability in a 20 pack-year smoker without a history of alcohol abuse.

None of the demographic variables (race, age, marital status) predicted relapse at 1 year, nor did current tobacco use measured in packs smoked per day. With the exception of the presence of CAD, no medical condition was associated with relapse.

Discussion

All the smokers enrolled in our study had the opportunity to turn an elective surgical procedure into a springboard for permanent tobacco cessation. Regulations at the San Francisco Veterans Affairs Medical Center proscribe smoking within the hospital; this information is conveyed to all patients at the time of hospitalization. While it is not uncommon for ambulatory patients to leave the hospital to smoke, it is not possible for patients to smoke on the day of their surgery and for several days after when confined to bed. Therefore, all SPI smokers quit using tobacco, if only for a brief period. Nonetheless, 76% of smokers had relapsed within a year. Relapse rates similar to ours have also been observed among patients hospitalized in coronary care units.[7]

The 1-year quit rate among SPI smokers, who did not have the benefit of an in-hospital smoking cessation intervention, was similar to the long-term quit rates associated with many outpatient stop smoking programs.[8,9] Analysis of the first National Health and Nutrition Examination Survey demonstrated an increase in the quit rate among smokers who had been hospitalized in the previous year.[10] Higher quit rates, up to 45%, have also been seen among smokers hospitalized with acute myocardial infarction.[11] Furthermore, our study demonstrates that the longer the hospitalization, probably reflecting poorer health status, the greater the likelihood of quitting.

Our study suggests that a history of addictive behavior, including long-term heavy tobacco use and alcohol abuse, may generally be predictive of smoking relapse. Other investigators have found that the measure "cigarettes smoked daily" does not predict those more likely to quit or less likely to relapse.[9,12] Pack-years of smoking, on the other hand, combines both duration with intensity of use, producing a measure of lifelong tobacco use that was clearly a better predictor of smoking relapse in our cohort than current tobacco use. Our patients with a history of alcohol abuse were also more likely to resume tobacco use. Alcohol is known to act as a relapse trigger along with other factors, such as stress, interpersonal conflict, and the presence of other smokers.[3]

There are a number of limitations to our study. First, our evaluation of predictors of smoking relapse was limited by the focus of the study—perioperative ischemia. We were therefore unable to assess the role of other potential correlates of smoking relapse, such as depression and education level. Thus, smoking cessation and relapse prevention programs for surgical patients will need to be based on a broader assessment of smoking relapse risk factors than can be provided from our study alone. Second, the SPI cohort was limited to men obtaining their surgical care at a Veterans Affairs hospital who had or were at high risk for CAD. Our results may therefore not be generalizable to younger, healthier patient populations with better social support systems and a lower risk of serious disease. Third, the pack-year history of smoking may have been underestimated if smokers reported cigarette consumption that was reduced prior to hospitalization. Fourth, all re-

ported quit rates were based on self-report by the study subjects. The actual relapse rate among the cohort is therefore likely to be somewhat greater than the 76% reported, particularly if the eight subjects lost to follow-up are also classified as relapsers. However, since smokers were enrolled in a study of the predictors of perioperative ischemia and not of smoking cessation, there was no incentive to overreport smoking cessation on follow-up. While nicotine intake is best assessed biochemically, self-report has been found to be a valid method for assessing smoking prevalence in observational studies.[3] Finally, because smoking relapse was common among our cohort, the ORs for pack-year history of smoking and history of alcohol abuse should not be interpreted as relative risks. The OR, a good approximation of the relative risk when an outcome is rare (generally less than 10%), overestimates the relative risk in our study because smoking relapse was so common (76%). The ORs nevertheless do reflect the importance of the association between the predictors in our study and smoking relapse.

The potential for achieving high long-term quit rates is demonstrated by the 71% quit rate achieved at 1 year among survivors of acute myocardial infarction who had the benefit of an in-hospital smoking cessation program.[11] This was a nurse-managed program using the principles of social learning theory and including outpatient follow-up. Similar programs focusing on smokers hospitalized for surgery—perhaps with special emphasis directed toward those most at risk for relapse, such as heavy long-term tobacco and alcohol users—could contribute to the continuing secular decline in tobacco use in the United States. Multimodality interventions that include self-help literature, audiovisual aids, nicotine gum, and counseling may prove the most effective. Well-designed randomized smoking cessation intervention trials for hospitalized patients are needed to answer this question. □

TABLE 2—Variables Associated with Smoking Relapse at 12 Months[a]

	OR (95% CI)[b]	P value
Univariate models		
History of alcohol abuse	2.69 (1.33, 5.40)	.006
Pack-year history (per 10 pack-years)	1.39 (1.05, 1.84)	.02
Postoperative days (per 7 days)	0.86 (0.75, 1.00)	.04
Coronary artery disease	0.54 (0.27, 1.07)	.08
	Adjusted OR (95% CI)[b]	**P Value**
Multivariate model		
History of alcohol abuse	3.07 (1.42, 6.66)	.005
Pack-year history (per 10 pack-years)	1.47 (1.09, 1.98)	.01
Postoperative days (per 7 days)	0.88 (0.75, 1.02)	.10
Coronary artery disease	0.49 (0.23, 1.03)	.06
Type of surgery[c]	0.50 (0.23, 1.10)	.09

[a]No other variables were significant at the $P < .10$ level.
[b]OR = odds ratio; CI = confidence interval. Because the risk of relapse was high, the ORs should not be interpreted as relative risks (see discussion).
[c]Major vascular and intrathoracic versus all other.

Acknowledgments

This research was supported by grant RO1-HL36744 from the National Institutes of Health.

The SPI (Study of Perioperative Ischemia) Research Group consists of the following individuals: *Principal investigator:* Dennis T. Mangano, PhD, MD; *associate investigators:* Martin J. London, MD, Warren S. Browner, MD, MPH, Milton Hollenberg, MD, Julio F. Tubau, MD, Jacqueline M. Leung, MD, William C. Krupski, MD, Joseph A. Rapp, MD, Marcus W. Hedgcock, MD, Edward D. Verrier, MD, Scott Merrick, MD; *technical staff/research associates:* M. Lou Meyer, MS, Linda Levenson, BS, Martin G. Wong, BS, RDMS, Elizabeth Layug, MD, Juliet Li, MD, Maria E. Franks, RN (former), Yuriko C. Wellington, MS (former), Mara Balasubramanian, MD (former), Evelyn Cembrano, MD (former), Wilfredo Velasco, MD, Safiullah N. Katiby, MD; *administrative/editorial assistance:* Thea Miller, *editorial advice:* Winifred von Ehrenburg; *research fellows:* Brian F. O'Kelly, MB, Jadwiga Szlachcic, MD, Andrew A. Knight, MD (former), Virginia Fegert MD (former), Paul Goehner, MD (former), David N. Harris, MD, FFARCS (former), Deanna Siliciano, MD, Nancy H. Mark, MD, Randy Smith, MD; *data management:* Ida M. Tateo, MS, Jeffrey Tice, BS, Cary Fox, MA, Angela Heithaus, BS (former); *consultants:* Jonathan Showstack, PhD, Diana C. Nicoll, MD, PhD, Paul Heineken, MD, Barry Massie, MD; *policy and data monitoring board:* Kanu Chatterjee, MB, FRCP, H. Barrie Fairley, MD, Lawrence W. Way, MD, Warren Winkelstein, MD, MPH (chairman).

References

1. Fiore MC, Novotny TE, Pierce JP, Hatziandreu EJ, Patel KM, Davis RM. Trends in cigarette smoking in the United States: the changing influence of gender and race. *JAMA.* 1989;261:49–55.
2. Orleans CT. Understanding and promoting smoking cessation: overview and guidelines for physician intervention. *Ann Rev Med.* 1985;36:51–61.
3. *Reducing the Health Consequences of Smoking: 25 Years of Progress: A Report of the Surgeon General.* Washington, DC: US Dept of Health and Human Services; 1989. Publication PHS 89–8411.
4. Health, United States, 1988. In: *Advance Data from Vital and Health Statistics.* Washington, DC: National Center for Health Statistics; 1989. DHHS publication PHS 89–1232.
5. Mangano DT, Browner WS, Hollenberg M, London MJ, Tubau JF, Tateo IM. Association of perioperative myocardial ischemia with cardiac morbidity and mortality in men undergoing noncardiac surgery. *N Engl J Med.* 1990;323:1781–1788.
6. SAS Version 5. Cary, NC: SAS Institute; 1985.
7. Rigotti NA, Singer DE, Mulley AG, Thibault GE. Smoking cessation following admission to a coronary care unit. *J Gen Intern Med.* 1991;6:305–311.
8. Hunt HA, Bespalec DA. An evaluation of current methods of modifying smoking behavior. *J Clin Psychol.* 1974;30:431–438.
9. Fiore MC, Novotny TE, Pierce JP, et al. Methods used to quit smoking in the United States. *JAMA.* 1990;263:2760–2765.
10. McWhorter WP, Boyd GM, Mattson ME. Predictors of quitting smoking: the NHANES I followup experience. *J Clin Epidemiol.* 1990;43:1399–1405.
11. Taylor CB, Houston-Miller N, Killen JD, DeBusk RF. Smoking cessation after acute myocardial infarction: effects of a nurse-managed intervention. *Ann Intern Med.* 1990;113:118–123.
12. Kabat GC, Wynder EL. Determinants of quitting smoking. *Am J Public Health.* 1987;77:1301–1305.

Smoking Cessation in Women Concerned about Weight

Phyllis L. Pirie, PhD, Colleen M. McBride, PhD, Wendy Hellerstedt, MPH, Robert W. Jeffery, PhD, Dorothy Hatsukami, PhD, Sharon Allen, MD, and Harry Lando, PhD

ABSTRACT

Background. **Weight gain after smoking cessation is often cited by women smokers as a primary reason for not attempting to quit smoking or for relapsing after a cessation attempt.**

Methods. **A randomized trial of 417 women smokers was conducted to test the addition of two weight control strategies to a smoking cessation program. Participants received the standard smoking cessation program, the program plus nicotine gum, the program plus behavioral weight control, or the program plus both nicotine gum and behavioral weight control. Weight and smoking status were measured at the end of treatment and at 6 and 12 months posttreatment.**

Results. **Smoking cessation rates were highest in the group receiving the smoking cessation program plus nicotine gum. Weight gain did not vary by treatment condition, so its effect on relapse could not be examined by group. There was no significant relationship between weight gained and relapse in individuals.**

Conclusions. **The added behavioral weight control program was attractive to the participants and did not reduce smoking cessation rates. However, it did not produce the expected effect on weight, thereby restricting our ability to examine the effect of weight control on smoking cessation and relapse. (*Am J Public Health.* 1992;82:1238–1243)**

Introduction

Men and women smokers in the United States are generally similar in their intentions to quit and their quit attempts,[1,2] and they are interested in similar types of cessation programs.[1] A major exception to this similarity is that women report substantially more concern about weight gain upon quitting[2] and are more likely than men to report using smoking as a weight control strategy.[3] Although it has been difficult to demonstrate that weight gain is a precipitant of relapse,[4] concern about weight gain may act as a barrier to quit attempts.

The research reported here is an investigation of treatments that incorporated weight control strategies into a well-tested smoking cessation program to create a program that is more acceptable and more attractive to weight-concerned women. The tested weight control strategies were a behavioral weight control program and nicotine polacrilex chewing gum. Nicotine gum, in addition to its demonstrated ability to enhance smoking cessation, has been reported to minimize weight gain among quitters in some studies[5–8] though not in others.[4,9]

Methods

Treatment Programs

The research design was a randomized factorial trial that tested nicotine gum and a behavioral weight control program alone and in combination; both were added components of the American Lung Association's Freedom from Smoking (FFS) clinic program. Participants were randomized to one of four groups: the FFS program; FFS plus the behavioral weight control program; FFS plus nicotine gum; or FFS plus both the behavioral weight control program and nicotine gum.

The FFS program, selected as one of the best widely available group programs,[10] is an 8-week program, with an orientation meeting plus seven treatment sessions. Meetings emphasize social support and the development of cognitive behavioral skills, such as anticipating and planning for high-risk situations, learning to cope, and preventing relapse. "Quit night" occurs in the fourth week of the program.

The behavioral weight control program was designed to help participants avoid weight gain and, should weight gain occur, to maintain a positive attitude about it. Weight control counseling accompanied each of the seven FFS sessions. Recommendations included decreasing caloric intake by 100 to 300 calories per day (depending on baseline cigarette consumption) and gradually increasing exercise to a goal of 1 hour of walking or its equivalent three times per week. Participants were given calorie counters and food diaries, which they were strongly encouraged to keep throughout the program.

Phyllis L. Pirie, Colleen M. McBride, Wendy Hellerstedt, Robert W. Jeffery, and Harry Lando are with the School of Public Health and Sharon Allen and Dorothy Hatsukami are with the School of Medicine at the University of Minnesota in Minneapolis.

Requests for reprints should be sent to Phyllis L. Pirie, PhD, Division of Epidemiology, School of Public Health, University of Minnesota, Minneapolis, MN 55455.

This paper was submitted to the Journal August 21, 1991, and accepted with revisions April 7, 1992.

Editor's Note. See related editorial by Giovino et al. (p 1203) in this issue.

In the groups assigned to receive nicotine chewing gum, the gum (2-mg strength) was supplied free of charge during treatment and for 3 months after treatment ended, and participants were strongly encouraged to use it. Specific training was provided in how best to use the gum, and time was allotted in each group session for discussion about and reinforcement of its use. No specific amount of gum was prescribed although participants were urged to chew at least three pieces a day.

Facilitators were female ex-smokers who had been previously trained by the American Lung Association to be group leaders for the FFS program. For this research project, additional training was provided regarding the specific additional treatments. Each facilitator was trained to administer all four treatment conditions, and facilitators were crossed by condition. On-site observers in all groups monitored adherence to the specified treatment condition to ensure that material was not introduced inappropriately.

Characteristics and Recruitment of Participants

Participants were women aged 20 to 64 who wanted to quit smoking and to maintain their weight while quitting; women who expressed a desire to gain weight were excluded. Exclusions for medical reasons (because nicotine gum would be contraindicated) included those who were pregnant or planning a pregnancy; those who were nursing an infant; those with a history of gastric ulcer, substance abuse in the past 6 months, or cardiovascular disease including hypertension; those with temporal mandibular joint syndrome; and those undergoing active cancer treatment.

Recruitment and all program activities were conducted primarily in a Minneapolis suburban community. Subjects were recruited from two sources. First, a previous research study involving a population-based screening for cardiovascular risk factors generated a list of 2631 age-eligible women smokers. These women were contacted by telephone for assessment of initial eligibility and were invited to participate. Upon initial screening, 1785 of them were found to be eligible for the study and 264 were eventually randomized into it. Second, because of initial concern that the list of smokers from the screening project would not yield a sufficient sample, additional volunteer participants were solicited by word of mouth and by two small newspaper advertisements. These sources yielded 249 interested women, of whom 153 were eventually randomized. Recruitment overall was more successful than had been projected, and the goal of recruiting and randomizing 360 individuals was surpassed by 57 individuals, for a total of 417 randomized subjects.

To facilitate the continuous recruitment and enrollment of subjects without delay in the beginning of treatment, treatment was administered in four consecutive cycles, with each treatment condition represented within each cycle. Thus, there was a total of 16 treatment groups.

Data Collection Procedures

Interested women were invited to attend a baseline clinic, which included the completion of demographic questionnaires, smoking history, and weight history; weight measurement with a balance beam scale; the assessment of possible contraindications to nicotine gum; and the collection of biochemical measures of smoking status (i.e., expired air for measuring carbon monoxide and saliva for measuring thiocyanate and cotinine). Participants at the baseline clinic were given a detailed explanation of the study and asked to sign a consent form indicating their acceptance of study procedures. Those signing the consent form were then randomly assigned to a weeknight and site for class participation. At the orientation meeting, the treatment assignment was announced to the group. All participants who attended this orientation meeting were considered participants in the full study.

Compliance with treatment recommendations was assessed by session attendance, by self-report items regarding completion of smoking-related "assignments" (e.g., smoking records, reading materials, self-assessment exercises, etc.), and by self-report items in the appropriate groups regarding completion of food records and use of nicotine gum. Weight was measured with a balance beam scale at each group session. Expired air carbon monoxide was measured 1 week after the quit date and at the end of treatment (4 weeks after the quit date). Saliva was collected for biochemical (both thiocyanate and cotinine) validation at the end of treatment and was analyzed in the laboratory only for those individuals reporting themselves to be abstinent.

Follow-up data collection was conducted in a clinic, with assessments similar to those conducted at baseline at 6 and 12 months after the end of treatment. A detailed questionnaire asked about current smoking status and smoking since the last contact, relapse situations, and withdrawal symptoms. As was done at the end of treatment, saliva was analyzed only for those individuals reporting themselves to be abstinent.

Monetary Contingencies

Participants were asked to make a $100 deposit at the time of the baseline clinic; this deposit was to be completely refunded upon completion of the 12-month follow-up clinic, regardless of smoking status at that time. A small number of participants who could not pay $100 paid smaller amounts but never less than $40.

Definitions

Two definitions of abstinence were used in this study: continuous abstinence (no smoking at all since the program quit night) and point prevalence abstinence (no smoking in the past 7 days). Self-reported abstinence was verified by an expired air carbon monoxide reading of less than or equal to 10 ppm; if the value was higher than 10 in a self-reported nonsmoker, we accepted a cotinine value of less than or equal to 10 ng/mL (in the groups not receiving nicotine gum) or a thiocyanate value of less than 100 μmol/L (in the groups receiving nicotine gum) as evidence of nonsmoking. According to these criteria, only one individual was believed to be misreporting smoking status at follow-up, and that person was treated as a smoker in the analyses.

Analysis Methods

Differences between treatment groups were examined at baseline, at the end of treatment (session 7 of the FFS program), and at 6 and 12 months posttreatment using χ^2 analysis for categorical variables and general linear model procedures for continuous variables. The SAS package of statistical programs was used.[11] Preliminary analyses revealed no design effects for smoking—that is, no differences within treatments within cycles that could be considered specific to individual cycles. Hence, the analysis of treatment conditions was based on all randomized individuals without accounting for the cycle in which they participated.

TABLE 1—Baseline Characteristics of Women's Smoking Project Participants, by Treatment Condition

	Treatment Condition							
	FFS (n = 103)		FFS and Nicotine Gum (n = 108)		FFS and Behavior (n = 109)		FFS, Nicotine Gum, and Behavior (n = 98)	
	Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI
Demographics								
Age, y	42.3	40.5, 44.1	42.9	41.1, 44.7	44.0	42.2, 45.8	43.4	41.6, 45.2
College graduates, %	31.3	22.3, 40.3	24.1	16.0, 32.2	27.8	19.4, 36.2	20.6	12.6, 28.6
Weight-related variables								
Weight, lb	140.0	136.3, 143.7	144.7	140.2, 149.2	144.6	140.1, 149.1	141.5	137.2, 145.8
Body mass index, kg/m^2	23.7	23.1, 24.3	24.2	23.4, 25.0	24.5	23.7, 25.3	24.0	23.4, 24.6
Expressing "great" weight concern, %	31.1	22.2, 40.0	41.7	32.4, 51.0	38.0	28.8, 47.2	31.6	22.4, 40.8
Smoking-related variables								
Cigarettes per day	25.6	23.2, 28.0	27.1	24.9, 29.3	26.9	24.7, 29.1	25.1	22.7, 27.5
Saliva cotinine	309.9	279.3, 340.5	274.3	244.3, 304.3	301.3	271.5, 331.1	284.0	252.4, 315.6
Smoking within 30 minutes of waking up in the morning, %	74.8	66.4, 83.2	80.6	73.1, 88.1	70.4	61.8, 79.0	76.5	68.1, 84.9

Note. FFS = Freedom from Smoking program; nicotine gum = nicotine gum condition; behavior = behavioral weight control condition.

Results

Baseline Comparisons

Table 1 shows selected baseline characteristics of participants randomized to each of the four treatment conditions. Participants in the four treatment conditions were statistically equivalent at baseline on all variables relevant to the treatment. Reported baseline smoking levels ranged from 4 to 105 cigarettes per day, with an average of 26.2 cigarettes. Although participants were concerned about weight, most did not appear to be obese; at baseline, they averaged 107% of the median recommended weight for height according to the recent (1983) life insurance industry tables.[12]

Process Measures

No incidents were recorded in which group leaders deviated from the assigned protocol for a given treatment group. Table 2 summarizes compliance data by treatment group. Meeting attendance was quite good, averaging 7.3 of 8 meetings. Self-reported compliance with treatment recommendations was more variable. Compliance with the smoking-related assignments differed across treatment groups in some sessions, with the best average for compliance reported in the group receiving both the behavioral weight control treatment and nicotine gum.

Follow-up Success

Data were collected in the follow-up clinics on 98.3% of participants at 6 months and on 98.1% of participants at 12 months. Those few participants who had moved out of the area were allowed to complete self-report assessments by telephone and mail, resulting in 100% participation in the follow-up assessments.

Weight Change Outcomes

Because weight gain was expected to be markedly different in smokers and nonsmokers and because length of abstinence was expected to be related to the amount of weight gained, weight change was analyzed at each follow-up point for the subgroup who remained continuously abstinent for the entire 12 months of follow-up (n = 88) and the subgroup who reported smoking continuously throughout the 12 months of follow-up (i.e., those who were smoking at each follow-up point) (n = 176). As expected, weight gain was pronounced in continuous abstainers, averaging 10.6 lb by the 12-month follow-up; it was very slight in the continuous smokers (1.6 lb by 12 months) (Table 3). Most of the weight gained was gained by the 6-month follow-up. Among those who were continuously abstinent, 70.5% gained at least 5 lb by the 12-month follow-up.

Although our treatment hypothesis was that abstinent individuals in groups receiving the behavioral weight control treatment or nicotine gum would demonstrate lower weight gains than individuals receiving the standard treatment, that was not the case (Table 4). Among the yearlong continuous abstainers, no significant differences in weight gain are seen between any of the treatment conditions at any follow-up point. During treatment, it appears that successful quitters in the treatment conditions that included the behavioral weight control treatment began to lose weight until quit night (session 3), after which their weights returned to and exceeded baseline levels (data not shown).

Among those abstinent at the end of treatment, weight gain was not found to be significantly different for those who were chewing gum versus those who were not, nor did it vary according to the amount of gum chewed. Similarly, at 6-month follow-up, weight gain was not found to be significantly different for those who were chewing gum versus those who were not.

Smoking Outcomes

Significant differences in abstinence among the treatment conditions are seen at each follow-up point (Table 5). The group receiving nicotine gum demonstrated the highest point prevalence and continuous abstinence rates at each follow-up point, whereas the group receiving just the basic smoking cessation program demonstrated the lowest smoking cessation rate at each follow-up point.

TABLE 2—Proportion of Participants Reporting Acceptable Levels of Compliance[a] with Treatment Recommendations, by Treatment Condition

	Treatment Condition											
	FFS			FFS, Nicotine Gum			FFS, Behavior			FFS, Nicotine Gum and Behavior		
	n	%	95% CI	n	%	95% CI	n	%	95% CI	n	%	95% CI
Attendance	103	92	87, 97	108	94	90, 99	108	86	80, 93	98	95	91, 99
Compliance with smoking program assignments												
Session 2	103	61	51, 71	108	69	61, 78	108	71	63, 80	98	77	69, 86
Session 3	103	25	17, 34	108	18	10, 25	108	26	18, 34	98	38	28, 47
Session 5	103	40	30, 49	108	29	20, 37	108	44	35, 54	98	55	45, 65
Session 6	103	60	51, 70	108	73	65, 82	108	56	46, 65	98	65	56, 75
Compliance with nicotine gum recommendation (of those not smoking; nicotine gum groups only)												
Session 5	...	...	...	81	97	94, 101	...	...	...	73	99	96, 101
Session 6	...	...	...	87	98	94, 101	...	...	...	67	97	93, 101
Session 7	...	...	...	76	93	88, 99	...	...	...	57	95	89, 101
Compliance with food records (behavioral treatment groups only)												
Session 2	...	...	...	...	...	...	108	69	61, 78	98	66	57, 76
Session 3	...	...	...	...	...	...	108	44	35, 54	98	55	45, 65
Session 5	...	...	...	...	...	...	108	29	20, 37	98	47	37, 57
Session 6	...	...	...	...	...	...	108	23	15, 31	98	32	22, 41

Note. FFS = Freedom from Smoking program; nicotine gum = nicotine gum condition; behavior = behavioral weight control condition.
[a]Acceptable levels of compliance are defined as follows: Attendance = present for at least seven sessions; compliance with smoking program assignments = self-report of at least "moderate" effort put into the assignment; compliance with nicotine gum recommendation = chewing at least three pieces per day; compliance with food records = self-report of keeping a record of "all" or "most" foods eaten. Those absent from a session are assumed to be below the acceptable level of compliance on measures taken at that session.

Given that we were unable to produce significant differences in weight gain across the treatment conditions, we are unable to analyze the effects of weight gain on smoking status at the group level. At the individual level, we found that, among successful quitters who had gained more than 5 lb by the end of treatment, the proportion who relapsed to smoking by the 6-month follow-up was 51.2% (95% CI = 36.3%, 66.1%), as compared with 39.9% (95% CI = 32.9%, 46.9%) in quitters who had gained 5 lb or less. Of those who were not smoking at the 6-month follow-up and who had gained more than 5 lb, the proportion who relapsed to smoking by the 12-month follow-up was 19.4% (95% CI = 11.8%, 27.0%). This compares with 13.2% (95% CI = 2.4%, 24.0%) relapsing among those who had gained less than 5 lb.

TABLE 3—Weight Gain (lb) from Orientation Meeting to End of Treatment and to 6- and 12-month Follow-ups, for Continuous Smokers and Continuous Abstainers, All Treatment Conditions Combined

	Continuous Smokers (n = 176)		Continuous Abstainers (n = 88)	
	Mean	95% CI	Mean	95% CI
By end of treatment	0.44	−0.21, 1.09	1.31	0.41, 2.21
By 6-month follow-up	1.56	0.44, 2.68	8.71	7.14, 10.29
By 12-month follow-up	1.62	0.35, 2.89	10.60	8.84, 12.4

Discussion

Self-reports from smokers, which suggest that weight gain is a common cause of relapse to cigarette smoking, lead to the expectation that the individuals who gained more weight following smoking cessation would demonstrate higher levels of relapse. Our ability to test this hypothesis was compromised, however, by our lack of success in producing differential weight changes among the treatment groups in this study. This outcome was unexpected, given that behavioral treatments of the type used here have been used successfully in many weight loss programs, and nicotine gum has often been shown to reduce weight gain in quitters who chew it. Compliance data may supply part of the explanation. Although excellent compliance was recorded for meeting attendance and use of nicotine gum, self-reported compliance with other treatment recommendations (e.g., completion of smoking-related assignments and of food records) was not as good. Compliance with the behavioral weight control intervention apparently declined following quit night, indicating perhaps that complying with two behavioral interventions simultaneously was difficult for participants.

At the individual level, women who

TABLE 4—Weight Gain (lb) from Orientation Meeting to End of Treatment and to 6- and 12-month Follow-ups, for Continuous Abstainers, by Treatment Condition

	Treatment Condition							
	FFS (n = 15)		FFS and Nicotine Gum (n = 34)		FFS and Behavior (n = 25)		FFS, Nicotine Gum, and Behavior (n = 14)	
	Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI
At end of treatment	2.42	0.40, 4.44	1.08	–0.27, 2.43	1.07	–0.50, 2.64	1.13	–0.97, 3.23
At 6-month follow-up	8.15	3.49, 12.81	7.96	4.86, 11.06	9.02	5.41, 12.63	10.61	5.79, 15.43
At 12-month follow-up	10.08	4.57, 15.59	9.97	6.30, 13.64	9.77	5.50, 14.04	14.15	8.45, 19.85

Note. FFS = Freedom from Smoking program; nicotine gum = nicotine gum condition; behavior = behavioral weight control condition.

TABLE 5—Smoking Abstinence at End of Treatment and at 6- and 12-Month Follow-ups, by Treatment Condition

	Treatment Condition							
	FFS (n = 103)		FFS and Nicotine Gum (n = 108)		FFS and Behavior (n = 108)		FFS, Nicotine Gum, and Behavior (n = 98)	
	%	95% CI	%	95% CI	%	95% CI	%	95% CI
Point prevalence abstinence[a]								
At end of treatment	45.6	36.0, 55.2	70.4	61.8, 79.0	53.7	44.3, 63.1	60.2	50.5, 69.9
At 6-month follow-up	22.3	14.3, 30.3	49.1	39.7, 58.5	31.5	22.7, 40.3	34.7	25.3, 44.1
At 12-month follow-up	19.4	11.8, 27.0	44.4	35.0, 53.8	27.8	19.4, 36.2	27.6	18.7, 36.5
Continuous abstinence[b]								
At end of treatment	37.9	28.5, 47.3	67.6	58.8, 76.4	44.4	35.0, 53.8	55.1	45.3, 64.9
At 6-month follow-up	16.5	9.3, 23.7	38.0	28.8, 47.2	29.6	21.0, 38.2	23.5	15.1, 31.9
At 12-month follow-up	14.6	7.8, 21.4	31.5	22.7, 40.3	23.1	15.2, 31.0	14.3	7.4, 21.2

Note. FFS = Freedom from Smoking program; nicotine gum = nicotine gum condition, behavior = behavioral weight control condition.
[a]No smoking in the past 7 days.
[b]No smoking at all since the program quit night.

gained more weight were not significantly more likely to relapse to smoking than women who gained less weight. At least one study[4] has found that those who gained more weight by the 6-month follow-up were more (rather than less) likely to be abstinent at 12 months. The results of the two studies may be partly related to the stage at which weight gain is assessed. It may be that extremely weight conscious individuals tend to respond to small increments in weight by relapsing in the early period. Among those who have managed to remain abstinent by 6 months, however, weight may be less of an issue. Hence, weight gain may be unrelated to further relapse, as in our study, or it may even be a marker for those who are strongly motivated enough to continue abstinence, as in the Hall et al. study.[4]

It is also surprising that we did not observe differential weight gain in those chewing and not chewing nicotine gum. Although many studies report the weight control effects of nicotine gum, it is not clear from the literature how much gum is needed to produce such effects, nor how long the gum must be used to produce this result. We saw no enhanced weight control, not even in those chewing 10 pieces a day at the end of treatment or in those continuing to use gum at 6 months as compared with those not using it.

As expected, this study demonstrated that adding nicotine gum to a group-based program can produce outstanding validated cessation results at the 12-month follow-up. This has also been demonstrated in other studies.[13] However, the treatment group receiving both nicotine gum and the behavioral weight control program as adjuncts to the FFS program produced smoking cessation rates that were significantly lower by the 12-month follow-up than those for the group receiving just nicotine gum as an adjunct to the basic program. The hypothesis that the combination of several treatments resulted in "overload" on the part of the participants is attractive but unproven. For example, there is no evidence that individuals in the group receiving the combined intervention demonstrated poorer compliance with treatment recommendations than those in the nicotine gum group; in fact, the group receiving the combined intervention generally had the highest compliance rates.

It is clear, though, that even the combined treatments did not detract from the smoking cessation outcome of the program. On average, those receiving the weight control treatments did as well as or better than those receiving the standard FFS program, and all the treatment conditions did about as well as or better than comparable group treatments reported elsewhere.[14] Thus, it is clearly not the case that adding weight control components will reduce the smoking cessation that can be achieved by a high-quality group program.

An unexpected observation made in the course of implementing this program was the attractiveness of the program to potential participants; the program roster was easily filled and even overfilled by the final enrollment cycle. In our previous ex-

periences with Minnesota communities, recruitment had been much more difficult. For example, Schmid et al.[15] had recruited about 1.4% of smokers in two communities to home correspondence courses, and Lando et al.[16] report that several communitywide smoking cessation contests had attracted 3% to 5% of smokers in each community. In contrast, this project, with careful eligibility screening and a required monetary deposit, succeeded in randomizing 14.8% of an identified population of smokers. It appears that smokers perceive a need for programs that will address their legitimate concerns about gaining weight when they quit smoking. It is possible that some of the participants in this program may not have been willing to participate in other, more traditional smoking cessation treatments. For this reason if for no other, continued investigation of combined weight control and smoking cessation treatments is warranted. □

Acknowledgments

This research was supported by grant no. CA41647 from the National Cancer Institute; Phyllis L. Pirie was principal investigator.

The assistance of Ms. Penny Gottier of the American Lung Association, Hennepin County affiliate, is gratefully acknowledged. The authors also wish to acknowledge the efforts of Ms. Karen Virnig, who organized and supervised the initial telephone contacts and much of the follow-up data collection, and Ms. Gail Dolliff, who organized and supervised the baseline and follow-up clinics.

References

1. Lando HA, Pirie PL, Hellerstedt WL, McGovern PG. Survey of smoking patterns, attitudes, and interest in quitting. *Am J Prev Med.* 1991;7:18–23.
2. Pirie PL, Murray DM, Luepker RB. Gender differences in cigarette smoking and quitting in a cohort of young adults. *Am J Public Health.* 1991;81:324–327.
3. Klesges RC, Klesges LM. Cigarette smoking as a dieting strategy in a university population. *Int J Eating Disord.* 1988;7:413–419.
4. Hall SM, Ginsberg D, Jones RT. Smoking cessation and weight gain. *J Consult Clin Psychol.* 1986;54:342–346.
5. Fagerstrom KO. Reducing the weight gain after stopping smoking. *Addict Behav.* 1987;12:91–93.
6. Emont SL, Cummings KM. Weight gain following smoking cessation: a possible role for nicotine replacement in weight management. *Addict Behav.* 1987;12:151–155.
7. Hajek P, Jackson P, Belcher M. Long-term use of nicotine chewing gum. *JAMA.* 1988; 260:1593–1596.
8. Gross J, Stitzer ML, Maldonado J. Nicotine replacement: effects on postcessation weight gain. *J Consult Clin Psychol.* 1989; 57:87–92.
9. Tonnesen P, Fryd V, Hansen M, et al. Two- and four-mg nicotine chewing gum and group counseling in smoking cessation: an open, randomized, controlled trial with a 22-month follow-up. *Addict Behav.* 1988; 13:17–27.
10. Lando HA, McGovern PG, Barrios F, Etringer BD. Comparative evaluation of American Cancer Society and American Lung Association smoking cessation clinics. *Am J Public Health.* 1990;80:554–559.
11. SAS Institute Inc. *SAS User's Guide.* Version 5. Cary, NC: SAS Institute Inc.; 1985: 956.
12. 1983 Metropolitan height and weight tables. *Stat Bull Metrop Insur Co.* 1984;64: 2–9.
13. Lam W, Sacks HS, Sze PC, Chalmers TC. Meta-analysis of randomized controlled trials of nicotine chewing gum. *Lancet.* 1987; ii:27–30.
14. Schwartz JL. *Review and Evaluation of Smoking Cessation Methods: The United States and Canada, 1978–1985.* Washington, DC: Division of Cancer Prevention and Control, National Cancer Institute; 1987:200. NIH publication 87–2940.
15. Schmid TL, Jeffery RW, Hellerstedt W. Direct mail recruitment to home-based smoking and weight control programs: a comparison of strategies. *Prev Med.* 1989; 18:503–517.
16. Lando HA, Loken B, Howard-Pitney B, Pechacek T. Community impact of a localized smoking cessation contest. *Am J Public Health.* 1990;80:601–603.

ASPO/Lamaze 1992 Childbirth Conference

From October 15 to 17, 1992, the American Society for Psychoprophylaxis in Obstetrics, Inc. (ASPO)/Lamaze will hold its annual conference at the Anaheim Hilton and Towers in Anaheim, Calif. The conference theme, "Childbirth: Choices and Challenges," reflects the diversity of issues and alternatives within the field of childbirth education. The highlights of the program will include dynamic keynote speakers, substantive educational sessions, dozens of exhibits, and special programs.

More than 600 childbirth educators, nurses, nurse-midwives, physicians, and consumers are expected to attend. Seminar topics will focus on the expanding roles, responsibilities, and technologies facing childbirth educators and parents today. Fifty-two professional development and educational sessions will be conducted by leading authorities in health care and Lamaze childbirth education. Optional programs include "Birth as a Cultural Mosaic," for both novice and expert educators, and "How to Establish a University ASPO/Lamaze Certification Program." The exhibit hall will feature more than 40 suppliers of products and services related to Lamaze childbirth education and perinatal health care.

For more information on the conference, contact ASPO/Lamaze at 1101 Connecticut Ave, NW, Suite 700, Washington, DC 20036; tel. (202) 857-1128.

How Women's Adopted Low-Fat Diets Affect Their Husbands

ABSTRACT

Objectives. One way of promoting a reduction in dietary fat intake is by changing the diet of family members. This study investigated the long-term effects of a low-fat dietary intervention on husbands of women who participated in the Women's Health Trial (WHT).

Methods. An average of 12 months after the end of the WHT, a randomly selected sample of participants' husbands was sent dietary and health questionnaires as part of a follow-up study of the maintenance of the low-fat diet among WHT participants.

Results. We found an absolute difference in fat intake between groups of 4 percentage points (32.9% energy from fat among intervention husbands [n = 188] vs 36.9% among control husbands [n = 180]). The wife's attitude and fat intake were among the most important predictors of her husband's fat intake, indicating that the effect of the WHT intervention on the husbands of participants was more likely due to their acceptance of lower-fat foods being served at home than to overt actions by the men.

Conclusions. Our results suggest that a dietary intervention aimed at women can have an effect on their husbands and may be a cost-effective approach to healthy dietary change for both women and men. (*Am J Public Health.* 1992;82:1244–1250)

Ann L. Shattuck, MPH, MS, RD, Emily White, PhD, and Alan R. Kristal, DrPH

Introduction

A diet high in fat has been implicated in both animal and epidemiologic studies as an important etiologic factor in a variety of diseases, including cardiovascular diseases and cancers of the colon, breast, and prostate.[1] *The Surgeon General's Report on Nutrition and Health* identified reduction of dietary fat consumption as the primary dietary priority for the general public as well as for special high-risk populations.[2] Recent research has focused on community-level interventions[3] and health promotion at worksites[4,5] and in schools[6,7] as potential effective and cost-efficient means of changing dietary and other health behaviors. In addition, the family has been identified as the primary unit in illness and health-related behaviors[8,9] and there is increased interest in effective ways to promote positive health habits within families.[6,10]

While there is evidence to suggest that eating patterns are shared within families,[11,12] there is little empirical research on the transfer of new dietary habits from one family member to another. Research has typically focused on family-based interventions[10,13,14] and on the effect of the family or spouse as a social support system on an individual's ability to modify dietary intake or other health behavior.[15] Only a few studies have addressed the impact of an intervention given to one individual on behavior change among his or her family members.[6,16–18] Although they suggest an important avenue of behavior change, these studies were small,[17] participants were self-selected or there was no control group,[6,18] or the research did not address the mechanism for the transfer of dietary patterns between spouses.[16] Thus, the purpose of this study was to address these shortcomings by studying the diffusion of dietary habits from women in the Women's Health Trial (WHT) to their husbands.

The WHT was designed to assess the feasibility of a large multicentered, randomized trial of a low-fat dietary intervention among women at moderately increased risk for breast cancer. Nutrient estimates based on a self-administered food frequency questionnaire indicated that intervention women had reduced the fat in their diet from 40% of energy intake at baseline to 26% at the end of the trial, compared with 39% and 38%, respectively, among control group women.[19] A follow-up study was conducted an average of 12 months after the end of the WHT to investigate maintenance of the low-fat dietary intervention and to determine predictors of this maintenance.[19–21] The intervention women had successfully maintained the low-fat diet at follow-up, with an (absolute) increase in dietary fat intake of only 1.4 percentage points.[19] At the time of the follow-up study, a randomly selected sample of husbands of WHT participants was also surveyed to assess the diffusion of the low-fat intervention from the participants to their spouses and to investigate the mechanism of this transfer. Specifically, the purpose of this study was to determine whether the husbands of intervention

The authors are with the Cancer Prevention Research Program, Fred Hutchinson Cancer Research Center, Seattle, Wash. Emily White and Alan R. Kristal are also with the Department of Epidemiology, University of Washington, Seattle.

Requests for reprints should be sent to Ann L. Shattuck, MPH, MS, RD, Fred Hutchinson Cancer Research Center, 1124 Columbia St, MP 702, Seattle, WA 98104.

This paper was submitted to the Journal October 25, 1991, and accepted with revisions April 7, 1992.

group women had reduced their dietary fat intake, compared with husbands of control group women, and to assess whether dietary change was due to the men's active pursuit of healthier eating habits or to their passive acceptance of their wives' new eating patterns.

Methods

The Women's Health Trial

The WHT began in three clinical sites in April 1985; however, despite successful recruitment and an effective intervention, the trial was canceled in March 1988 because of concerns about the costs and benefits of a large low-fat intervention trial. The details of the first phase of the trial and of the intensive low-fat dietary intervention have been reported elsewhere.[22,23] Briefly, the WHT participants were healthy women aged 45 to 69 years who, after being screened for eligibility, were randomly assigned to the dietary intervention group or the control group. The intervention women attended small-group sessions, which met weekly and then monthly for a period of 5 to 37 (mean of 15.5) months. The major goal was to reduce total dietary fat intake by intervention women to 20% of baseline kilocalories by means of nutrition education and behavior modification strategies. The objective in the control group was to intervene as little as possible with the participants' normal diet while collecting follow-up medical and dietary data.

Husbands of participants were not formally included in the intervention and very few attended educational sessions, but all intervention husbands were invited to attend infrequent (e.g., annual) group activities. Intervention women were encouraged to cook low-fat meals for their families, but they were also provided with methods for incorporating their low-fat diet as a subset of the regular family meals.

Sample Selection

WHT intervention women and their husbands were eligible for the follow-up study if the women had begun intervention sessions before November 1, 1987. Control group women and their husbands were eligible if the women had been randomized before November 1, 1987. The cutoff date allowed intervention women to complete a minimum of eight dietary education sessions before the announcement of the end of the trial. Women who were diagnosed with breast cancer during the trial (and their husbands) were not eligible for the follow-up study. From 1650 eligible WHT women, 525 intervention women and 525 control women were randomized into the follow-up study; of these, 793 were married at baseline. Two hundred twenty-five intervention husbands and 225 control husbands, divided equally among the three clinical sites, were randomly selected for this study. Men were eligible to participate only if they were married to and living with a WHT participant at baseline and at follow-up.

Data Collection

Data collection was conducted between January and December 1989, and was randomly distributed over a period of 6 to 20 (average of 12) months after active WHT intervention had ended. In addition, data collection was equally distributed across months among intervention and control groups. Dietary data, food patterns, health information, and demographic data were collected by means of mailed, self-administered questionnaires. Twelve weeks after the initial mailing, telephone interviews were attempted for all participants who had not yet refused or responded to the previous two mailings and reminder telephone call.

A 100-item semiquantitative food frequency questionnaire was completed by the women and their husbands and was used to estimate average daily intake of energy, fat, and other nutrients. The questionnaire, nutrient data base, and computer software program for analysis of the questionnaire were modifications of those developed by Block and colleagues at the National Cancer Institute (NCI).[24,25] Foods such as frozen yogurt and low-fat cheeses were added to the food frequency questionnaire to enhance the assessment of low-fat food selection. In addition, to reflect current dietary patterns, the reference time period for the food frequency questionnaire was "over the last 3 months."

For the analyses presented here, several of the items from the husbands' health questionnaire were combined into scales. An indicator of socioeconomic status was computed, on the basis of Hollingshead's two-factor index of social position,[26] by combining scales for self-reported education and occupation. A knowledge score was calculated by summing the number of correct answers to whether or not 14 food items were high in fat. The attitude score was the average of the responses (scored from 1, "disagree strongly," to 5, "agree strongly") to two items on diet and chronic disease: "A low-fat diet can reduce your risk of heart disease" and "A low-fat diet can reduce your risk of cancer." A summary score representing positive changes in eating patterns was created by summing the number of high-fat foods that a husband reported eating less often and the number of low-fat foods that he reported eating more often. A higher dietary change score (range 0 to 12) indicated a greater number of fat-reducing dietary changes.

Response Rates and Missing Data

Four hundred thirty-six of the 450 randomized husbands were eligible to participate in the follow-up study: 6 intervention and 7 control husbands were excluded because the wives reported they were no longer married (i.e., that they were separated, divorced, or widowed); 1 control husband was excluded because his wife had developed breast cancer after the end of the WHT and dietary changes might have followed the diagnosis. Seventy-one percent (156/219) of the eligible intervention husbands and 68% (148/217) of the eligible control husbands completed the mailed questionnaires. Of those who did not complete the mailed questionnaires, about half of the intervention husbands (n = 35) and the control husbands (n = 32) completed a shortened telephone version of the health questionnaire that included questions on weight and eating patterns.

Statistical Analysis

Differences between intervention and control group husbands were examined with Student's t test for continuous variables, with a chi-square test for categorical variables, and with the Mantel-Haenszel test for trend for ordered categorical variables.[27] With the exception of nutrient density values, dietary variables were log transformed to normalize their distribution and then untransformed for presentation in the tables. Regression analyses with percentage of energy from fat as the dependent variable were performed to adjust for potential confounding factors. Regression models were also used to determine predictors of a lower-fat diet while controlling for an intervention-group effect, and for effect modification by intervention group (interaction between each factor and group).[28] We explicitly tested factors that we hypothesized to be intermediate between the wife's participation in the intervention (group effect) and any resulting change in the husband's dietary fat intake by fitting regression mod-

TABLE 1—Demographic and Health Characteristics of the Husbands of Intervention and Control Participants in the Women's Health Trial

	Intervention (n = 156)		Control (n = 148)		
	No.	%	No.	%	*P* Value for Difference[a]
Age, y					
30–49	11	7.1	4	2.7	
50–59	59	38.1	73	49.3	
60–69	74	47.7	59	39.9	
70+	11	7.1	12	8.1	.984[b]
Current smoker	10	6.7	23	15.6	.013
Exercise					
Rarely or never	41	27.0	55	37.4	
A few times per month	43	28.3	33	22.5	
1–2 times per week	33	21.7	21	14.3	
3+ times per week	35	23.0	38	25.9	.108
Self-rating of health					
Excellent	43	28.1	34	23.1	
Better than average	57	37.3	69	46.9	
Average	51	33.3	36	24.5	
Worse than average	2	1.3	8	5.4	.642[b]
History of disease					
Cancer	12	8.0	10	6.9	.731
Heart disease/ angina	15	9.9	13	9.0	.776
Heart attack	8	5.3	9	6.2	.726
High blood cholesterol	44	29.1	38	26.0	.549
Hypertension	48	31.4	49	33.8	.656
Stroke	6	4.0	2	1.4	.284[c]
Diabetes	9	5.9	10	6.9	.731
Current working status					
Full-time	89	58.9	89	60.5	
Part-time	18	11.9	17	11.6	
Retired	44	29.1	41	27.9	.960
	Mean	SD	Mean	SD	*P* Value for Difference[d]
Education, y	15.7	3.2	15.5	3.0	.568
Socioeconomic status[e]	2.2	1.1	2.4	1.0	.151

[a] *P* values are based on chi-square test unless otherwise noted.
[b] Mantel-Haenszel test for trend.
[c] Fisher's Exact Test (two-tail).
[d] Student's *t* test.
[e] Socioeconomic status is denoted by Hollingshead's two-factor index of social position, in which 1 = highest, 5 = lowest.

els to predict husbands' fat intake based on each factor plus the group effect (see Table 5, note a).

Results

The husbands of the two groups of WHT participants (intervention and control) were similar in terms of reported demographic characteristics and health status at follow-up (Table 1). Most of the men were between 50 and 69 years of age, well-educated, healthy, above the middle level of Hollingshead's 5-point scale of socioeconomic status,[26] and working at least part-time. Significantly fewer intervention husbands than control husbands smoked cigarettes at follow-up ($P < .01$). There were no differences between groups in the number of months the wives participated in WHT (a mean of 15 months) or in the time between the end of the trial and follow-up (a mean of 12 months) (data not shown).

The husbands' mean daily intake of energy, fat, and other major nutrients, as estimated from a self-administered food frequency questionnaire, are summarized in Table 2. An average of 1 year after the trial ended, intervention husbands consumed significantly fewer kilocalories and less dietary fat and cholesterol than did control husbands (all Ps $< .01$). As a percentage of energy intake (%en), total dietary fat intake was 32.9% for intervention husbands vs 36.9% for control husbands ($P < .001$). This difference was similar across the three geographic locations and did not change when adjusted for smoking habits or history of chronic diseases (data not shown). There were no differences between groups in dietary fiber consumption or in the ratio of polyunsaturated (linoleic) to saturated fatty acids.

The husbands were asked about how their consumption of selected foods and food groups had changed since their wives entered the WHT (Table 3). Significantly more intervention than control husbands reported eating less of each high-fat food (all Ps $< .01$), but there were no significant differences between groups in increased consumption of foods promoted as healthy. The decreased consumption of foods high in both saturated and polyunsaturated fat is consistent with the WHT goal of lowering total fat. The summary dietary change score differed substantially between intervention and control group husbands ($P < .001$), indicating that intervention husbands made significantly more dietary changes in the recommended direction (e.g., decreasing high-fat foods) than did the control husbands.

Mean recalled body weight at the time their wives entered WHT did not differ between intervention and control group husbands (Table 4). At follow-up, however, intervention husbands had a significantly lower mean reported weight than did control husbands ($P < .05$). Consistent with the lower estimated energy intake by the intervention husbands, weight change between the two time periods was greater in the intervention group than in the control group ($P = .06$).

To understand whether the reduction in fat intake among intervention husbands was due to an active change in the husbands' knowledge, attitudes, and behaviors or to a passive acceptance of low-fat meals consumed at home, we compared measures of these factors between groups (Table 5). In addition, we tested these variables as intermediate factors between the WHT intervention and a change in the husbands' fat intake; an "intermediate" would be expected to significantly predict husbands' fat intake (β_1) and to reduce the group effect (β_2) (from a coefficient of 3.96 in a model with group alone). Lastly, we anticipated that some of these variables would have a greater effect on dietary fat intake among intervention husbands than among control group husbands; however,

there was no evidence of effect modification for any of the measured factors (data not shown).

Although knowledge of foods high in fat was significantly greater among intervention husbands than among control group husbands, knowledge did not have an effect on overall fat intake, nor did it explain the group effect (Table 5). Conversely, the husbands' attitude toward diet and disease was strongly associated with their percent of energy from fat, but there was no difference in attitude between intervention and control group husbands. The number of meals that a husband and wife shared each week had a significant, independent effect on the husband's fat intake; however, the number of meals shared did not reduce the intervention effect. There were no differences between the intervention and control groups in the reported frequency with which husbands chose low-fat foods when not eating with their wives. Although it was strongly associated with husbands' percentage of energy from fat across both groups of men, after we adjusted for intervention group this variable made no difference in the coefficient for group effect (β_2).

As expected, a husband's fat intake was positively associated with his wife's percentage of energy from fat ($P < .001$). A correlation of 0.34 was found between a husband's percentage of energy from fat and his wife's ($P < .001$, data not shown); this was a high correlation, considering the degree of measurement error in both variables. Of all the factors listed in Table 5, only the wife's fat intake substantially explained the intervention group effect (β_2 was reduced and no longer significant). In addition, the husbands of intervention group women who reported fixing separate meals for their families consumed a diet 3.6 (absolute) percentage points higher in fat (as %en) than intervention group husbands whose wives did not prepare separate meals ($P < .01$).

TABLE 2—Mean Daily Nutrient Intake by the Husbands of Women's Health Trial Participants Approximately 1 Year after the Low-Fat Intervention Trial Ended

	Intervention (n = 156)		Control (n = 148)		*P* Value for Difference[a]
	Mean	SD	Mean	SD	
Nutrient[b]					
Energy, kcal	1551.6	574.7	1736.6	654.7	.009
Fat, g	54.8	28.1	69.5	34.7	.001
Saturated	19.0	10.1	23.9	12.8	.001
Polyunsaturated[c]	19.7	10.5	25.2	13.1	.001
Monounsaturated[d]	9.6	6.1	13.1	7.4	.001
Protein, g	64.1	24.9	68.8	25.8	.102
Carbohydrate, g	172.5	68.8	185.5	69.0	.101
Dietary fiber, g	12.9	5.5	13.4	5.4	.370
Cholesterol, mg	199.0	108.2	237.0	134.3	.006
Nutrient density					
Fat, % kcal	32.9	8.2	36.9	7.9	.001
Protein, % kcal	16.8	2.9	16.2	3.4	.125
Carbohydrate, % kcal	45.4	8.8	43.4	7.8	.041
Cholesterol, mg/1000 kcal	137.5	56.3	149.5	81.9	.139
Dietary fiber, g/1000 kcal	8.9	3.3	8.3	3.1	.141
Ratio of polyunsaturated to saturated fatty acids	0.54	0.20	0.58	0.22	.151

[a]*P* values for unequal variances are presented where appropriate.
[b]Nutrients were analyzed with natural log transformations and then untransformed, yielding the geometric mean [exp(z), where $z = \ln(\text{nutrient})$] and standard deviation [SD = exp(z)SD(z)].
[c]Linoleic acid only.
[d]Oleic acid only.

Discussion

We assessed the transfer of dietary habits from wives to husbands 1 year after the wives' participation in an intensive low-fat dietary intervention had ended. The intake of total dietary fat and of each type of fat was significantly lower among the husbands of the intervention group women than among the husbands of control group women. Nutrient estimates from a self-administered food frequency questionnaire indicated that intervention husbands consumed 32.9%en as fat, vs 36.9%en for control husbands. The mean fat intake by control group men is in close agreement with data from the National Health and Nutrition Examination Survey II (1976 to 1980),[29] which found a fat intake of approximately 37%en for men over 45 years of age, and with the 38%en for men aged 18 to 65 years reported by Stephen and Wald[30] in a combined analysis of 171 studies that obtained individual dietary data (1980 to 1985). Thus, the dietary fat intake (as %en) of the participants in this study appears to be similar to that reported elsewhere for representative men in this country.

Although the 4-percentage-point group difference in percentage of energy from fat (32.9% vs 36.9% for intervention and control husbands, respectively) is only half the 8-point difference observed in the WHT women at follow-up (27.7% vs 35.1%),[19] it represents considerable progress toward the 30%en from total fat recommended by the NCI and other national health organizations. Furthermore, the magnitude of the group difference is similar to that reported by some large-scale dietary trials that intervened directly with high-risk men.[31,32] For example, the Multiple Risk Factor Intervention Trial (MRFIT) reported 34%en as fat in the special intervention group (meeting the study goal) vs 38%en in the usual-care group; these figures were mean percentages of energy from fat averaged from follow-up data for years 1, 3, and 6.[31]

The group difference in fat intake found in the present study is twice the 2-percentage-point difference (34%en vs 36%en) found between intervention and control husbands during the WHT.[16] Several reasons for the larger difference in the follow-up study have been considered. Different instruments were used to measure dietary intake, and the current version (Block, NCI)[24] was modified to be more sensitive to lower fat intake. Greater measurement error in the dietary assessment made during the trial may have masked a larger effect. In addition, acceptance of low-fat dietary patterns by the husbands may occur slowly, or sufficient knowledge of the diet to accurately complete a food questionnaire may be acquired over time; therefore, greater differences would be expected after an extended time period.

The changes in food consumption patterns among the husbands in this study reflect the WHT intervention[23,33] more than the healthy eating habits recommended for general health and heart dis-

TABLE 3—Reported Change in Consumption of Selected Foods by Women's Health Trial Intervention and Control Husbands at Follow-up

	Intervention		Control		
	n[a]	Eat Less, %	n[a]	Eat Less, %	*P* Value[b]
High-fat foods					
Whole milk	112	67.0	110	43.6	.001
Butter	128	78.1	121	52.9	.001
Margarine	172	37.2	173	17.3	.001[c]
Eggs	176	77.8	174	54.6	.001
Cheese	182	52.8	174	25.9	.001
Red meat	184	71.7	178	56.7	.005
Cakes, cookies, ice cream	177	48.6	177	32.8	.002[c]
	n[a]	Eat More, %	n[a]	Eat More, %	*P* Value[b]
Other foods					
Fish	180	62.8	176	55.1	.228
Fruits	185	46.5	178	42.1	.398
Vegetables	185	41.1	178	32.0	.101
Whole wheat breads	166	45.2	165	35.2	.165[c]
Whole grains	161	47.8	164	40.2	.329[c]
	n Mean SD		n Mean SD		
Total change score[d]	185 6.0 2.8		180 4.4 2.8		.001[e]

[a]Number of men who consumed the foods.
[b]The intervention–control group difference was based on Fisher's Exact Test (two-tail) unless otherwise noted.
[c]Chi-square test.
[d]The score range was 1–12, increasing with number of dietary changes made in the recommended direction (e.g., lower fat, higher fiber).
[e]Student's *t* test.

TABLE 4—Self-reported Mean Body Weight and Weight Change of Husbands of Women's Health Trial Participants

	Intervention (n = 185)		Control (n = 178)		
Weight, lb	Mean	SD	Mean	SD	*P* Value for Difference
Baseline	187.3	29.9	192.1	31.2	.136
Follow-up	185.3	29.0	191.8	32.0	.044
Weight change	−2.1	9.0	−0.4	8.3	.063

ease prevention that have been widely publicized over the past few decades. While all of the foods listed in Table 3 represent foods for which the NCI, the American Heart Association, and other health organizations have recommended changes,[1] the foods for which consumption was significantly different between intervention and control husbands were ones associated with lowering total fat (saturated and unsaturated fatty acids) in the diet. Correctly identifying foods high in polyunsaturated fat (peanut butter and margarine) was also more common among intervention than control men, suggesting some transfer of knowledge from intervention wives concerning a diet low in total fat. There were, however, no differences between groups in other promoted healthy eating habits, such as increased intake of high-fiber fruits, vegetables, and whole grains. This increases our confidence that the changes in the dietary habits of intervention group husbands were attributable to the WHT low-fat intervention.

Our finding that an intervention given to one family member can have a positive effect on the health behaviors of others supports the results of several other studies.[6,17,18] In studies relevant to the design of this study, Zimmerman et al.[17] and Sexton et al.[18] studied wives of men who participated in intervention trials. Zimmerman and colleagues[17] compared the effects of a work-site health promotion program for fire fighters on 32 intervention group wives compared with 9 wives of fire fighters who had not yet received the program. While no statistically significant differences in the wives' health habits were found, healthy dietary changes were made in the expected direction. Sexton et al.[18] assessed risk factor changes in 220 wives of MRFIT special-intervention (SI, n = 80) and usual-care (UC, n = 140) participants. Beneficial effects of the MRFIT intervention were suggested by lower food-record ratings (indicating greater adherence to the MRFIT eating plan), higher diet-knowledge scores, and a lower mean weight and serum cholesterol among the SI wives. However, only 36% of the invited SI wives and 64% of the UC wives participated in the study, and self-selection may have led to greater between-group differences than would otherwise be the case. The results of these studies and our research, however, do suggest the potential importance of the effects on other family members of one member's participation in a health-oriented intervention program.

The previous report on WHT husbands[16] made no direct attempt to investigate whether the lower fat intake by intervention group husbands was due to their passive acceptance of low-fat foods being served by their wives or to the husbands' active participation and internalization of their wives' new eating patterns. In the current study we were able to address this question more directly, and a number of our findings suggest that the former case is more likely (Table 5). First, there was no difference between the two groups of husbands in their responses as to whether they chose low-fat foods when not eating with their wives. Second, our hypothesized intermediate variables, knowledge of foods high in fat and attitude regarding diet and disease prevention, did not appear to explain the intervention effect. Finally, the wife's attitude and behavior were among the most important predictors of a husband's dietary fat intake. This is consistent with findings by Schafer,[34] who reported that a wife's influence, and a husband's perception of his wife's influence, were related to better diets for both husbands and wives. The women in our study had the primary responsibility for family meal preparation: over 70% of the women in both groups reported having at least 90% of the responsibility for planning, purchasing, and preparing family meals. In addition, we found

TABLE 5—Knowledge, Attitudes, and Dietary Behavior of Women's Health Trial Participants and Husbands, as Predictors of Husbands' Fat Intake

	Comparison of Intervention and Control Groups							Predictor of Husbands' Fat Intake[a]			
	Intervention			Control			*P* Value for Difference				
	No.	Mean	SD	No.	Mean	SD		Factor β1	*P* Value	Group β2	*P* Value
Husbands' attitudes and behavior											
Knowledge score[b]	153	11.9	1.9	147	11.1	2.2	.001	−0.26	.255	3.81	.001
Attitude score[c]	153	4.4	0.6	147	4.3	0.7	.176	−3.54	.001	3.89	.001
No. meals per week eat same foods as wife	187	10.4	4.8	180	10.2	4.5	.658	−0.23	.023	3.97	.001
Choose low-fat foods when not eating with wife[d]	179	2.7	0.9	171	2.6	0.9	.313	−3.38	.001	3.99	.001
Wives' attitudes and behavior											
Wife's fat intake, %en	166	27.7	7.9	166	36.3	8.2	.001	0.26	.001	1.81	.062
Fixed separate meals[e]	161	1.4	0.5					3.63	.008		
Husband encouraged diet[e]	151	1.6	0.5					−0.28	.842		
Wife's diet good example[e]	161	1.9	0.3					−2.58	.204		

[a]Regression model: Percentage of energy intake (%en) from fat = $\alpha + \beta_1$ factor + β_2 group. Limited to the 304 participants who completed the food frequency questionnaire.
[b]Based on total number of right answers to whether or not 14 food items were high in fat.
[c]The average of responses to two items on diet and health; a high score is consistent with the belief that diet improves health. Scale: 1 = disagree strongly, 2 = disagree somewhat, 3 = don't know, 4 = agree somewhat, 5 = agree strongly.
[d]Scale: 1 = rarely or never, 2 = sometimes, 3 = often, 4 = usually or always.
[e]As reported by wife, intervention women only; coded 1 = No, 2 = Yes.

a high correlation between a husband's percentage of energy from fat and his wife's. Similarly, Lee and Kolonel found high correlations in nutrient intake among 1428 married couples[11] and good correspondence for food items most likely to be consumed at home among another sample of 281 husband-wife pairs.[12] In summary, the effect of the WHT low-fat dietary intervention on husbands seemed to be accomplished through the behavior of their wives. The above findings taken together suggest that the lower fat intake reported by the intervention group husbands was more likely due to their passive acceptance of lower-fat foods being served at home than to their active participation in their wives' new eating patterns.

A number of limitations need to be considered in interpreting the results of this study. No data were obtained from the husbands at the time their wives entered the WHT. However, because of the randomization of women into intervention and control groups and the random selection of women and their husbands for this follow-up study, the baseline characteristics of the intervention and control group husbands were assumed to be comparable. With the exception of smoking, there were no differences in the husbands' self-reported demographic or health characteristics at the time of this study, and controlling for smoking in a multiple regression model did not change the between-group difference in percentage of energy from fat. If we are correct in assuming baseline similarities between groups, a greater reduction in dietary fat intake occurred among intervention husbands than among control husbands.

Social desirability bias could have influenced intervention group husbands to report lower fat intakes. We also recognize that bias may be present in the husbands' self-reported dietary changes, especially as they relate directly to the WHT intervention. However, because the husbands were not active participants in the WHT dietary intervention, they would be less likely than their wives to be influenced by these biases.

There is also a potential for bias due to the self-selection of husbands who completed the study questionnaires. Seventy percent of eligible intervention husbands completed the mailed questionnaires, and half of the 30% nonrespondents (35 intervention and 32 control husbands) completed a shorter questionnaire by telephone. To evaluate the potential selection bias, we compared the dietary change scores (see Table 3) of mail respondents with those of telephone respondents. The magnitude of the difference between mail and telephone respondents was identical for intervention and control groups, indicating that there was no differential bias between groups.

The women who participated in the WHT were a group of dedicated and highly motivated women, which limits the generalizability of our findings. The changes in dietary fat intake made[22,23] and maintained by the intervention women[19] may be greater than could be expected in a less select population. Therefore, the magnitude of change seen in the husbands' diets may also be greater than could be expected from a less intensive intervention. On the other hand, the observed difference in fat intake between the two groups may be less than the actual difference. In the follow-up study, we found that the low-fat dietary behaviors most easily maintained by the WHT intervention women were the substitution of specially manufactured low-fat foods and the modification of commonly used foods to be lower in fat.[20] Because these dietary changes are not likely to be noticed by most husbands, the changes in dietary fat intake may be underreported, indicating a potentially larger intervention group–control group difference in percentage of energy from fat than was measured. Furthermore, the group difference may also have been underestimated because of the measurement error inherent in the food frequency instrument and because of the instrument's insensitivity to low fat intake.[19,33]

We conclude that the WHT dietary intervention program for motivated women had a significant effect on nonparticipating spouses at no additional cost. The mechanism of this effect appeared to be passive change brought about by the

success of the WHT intervention and the husbands' acceptance of low-fat foods being served at home. The 34%en from fat consumed by intervention husbands is a substantial move toward the dietary goal of 30%en recommended by many US health organizations. Although this was an intensive intervention with a select group of motivated women, an intervention directed at, or including, the family member with food purchasing and preparation responsibilities could be a cost-effective and broad-reaching approach to public health dietary change programs. □

Acknowledgments

This research was supported by Grant CA 34847 from the National Cancer Institute. Support for the Women's Health Trial came from National Cancer Institute Cooperative Agreements CA 38413, CA 38511, CA 38526, CA 38550, and CA 38331.

The authors wish to acknowledge the principal investigators for the Women's Health Trial: Ross L. Prentice, PhD, Statistical Coordinating Center; Maureen M. Henderson, MD, Seattle Clinical Center; Sherwood L. Gorbach, MD, Nutrition Coordinating Center; William Insull, Jr, MD, Houston Clinical Center; Myron Moskowitz, MD, Cincinnati Clinical Center.

References

1. Committee on Diet and Health, National Research Council. *Diet and Health: Implications for Reducing Chronic Disease Risk.* Washington, DC: National Academy Press; 1989.
2. *The Surgeon General's Report on Nutrition and Health.* Washington, DC: US Dept of Health and Human Services; 1988: 9–11. DHHS publication PHS 88-50210.
3. Matarazzo JD, Weiss SM, Herd JA, Miller NE, Weiss SM, eds. *Behavioral Health: A Handbook of Health Enhancement and Disease Prevention.* New York, NY: John Wiley & Sons Inc; 1984:1154–1178.
4. Masur-Levy P, Tavris DR, Elsey-Pica L. Cardiovascular risk changes in a work-site health promotion program. *J Am Diet Assoc.* 1990;90:1427–1428.
5. Maysey DL, Gimarc JD, Kronenfeld JJ. School worksite wellness programs: a strategy for achieving the 1990 goals for a healthier America. *Health Educ Q.* 1988; 15:53–62.
6. Coates TJ, Jeffery RW, Slinkard LA. Heart healthy eating and exercise: introducing and maintaining changes in health behaviors. *Am J Public Health.* 1981;71:15–23.
7. Nader PR, Perry C, Maccoby N, et al. Adolescent perceptions of family health behavior: a tenth grade educational activity to increase family awareness of a community cardiovascular risk reduction program. *J School Health.* 1982:372–377.
8. Litman TJ. The family as a basic unit in health and medical care: a social-behavioral overview. *Soc Sci Med.* 1974;8:495–519.
9. Osterweis M, Bush PJ, Zuckerman AE. Family context as a predictor of individual medicine use. *Soc Sci Med.* 1979;13A:287–291.
10. Nader PR, Sallis JF, Patterson TL, et al. A family approach to cardiovascular risk reduction: results from the San Diego Family Health Project. *Health Educ Q.* 1989;16: 229–244.
11. Lee J, Kolonel LN. Nutrient intakes of husbands and wives: implications for epidemiologic research. *Am J Epidemiol.* 1982;115:515–525.
12. Kolonel LN, Lee J. Husband-wife correspondence in smoking, drinking, and dietary habits. *Am J Clin Nutr.* 1981;34:99–104.
13. Knutsen SF, Knutsen R. The Tromso Heart Study: family approach to intervention on CHD. *Scand J Soc Med.* 1989;17: 109–119.
14. Carmody TP, Istvan J, Matarazzo JD, et al. Applications of social learning theory in the promotion of heart-healthy diets: the Family Heart Study intervention model. *Health Educ Res.* 1986;1:13–27.
15. Coppotelli HC, Orleans CT. Partner support and other determinants of smoking cessation maintenance among women. *J Consult Clin Psychol.* 1985;53:455–460.
16. White E, Hurlich M, Thompson RS, et al. Dietary changes among husbands of participants in a low-fat dietary intervention. *Am J Prev Med.* 1991;7:319–325.
17. Zimmerman RS, Gerace TA, Smith JC, Benezra J. The effects of a worksite health promotion program on the wives of fire fighters. *Soc Sci Med.* 1988;26:537–543.
18. Sexton M, Bross D, Hebel JR, et al. Risk-factor changes in wives with husbands at high risk of coronary heart disease (CHD): the spin-off effect. *J Behav Med.* 1987;10: 251–261.
19. White E, Shattuck AL, Kristal AR, et al. Maintenance of a low-fat diet: follow-up of the Women's Health Trial. *CA Epidemiol, Biomarkers Prev.* 1992;1:315–323.
20. Kristal AR, White E, Shattuck AL, et al. Long-term maintenance of low-fat diets: durability of fat-related behavior in the Women's Health Trial. *J Am Diet Assoc.* 1992; 92:553–559.
21. Urban N, White E, Anderson G, Curry S, Kristal A. Correlates of maintenance of a low-fat diet among women in the Women's Health Trial. *Prev Med.* 1992;21:279–291.
22. Henderson MM, Kushi LH, Thompson DJ, et al. Feasibility of a randomized trial of a low-fat diet for the prevention of breast cancer: dietary compliance in the Women's Health Trial Vanguard Study. *Prev Med.* 1990;19:115–133.
23. Insull W, Henderson MM, Prentice RL, et al. Results of a randomized feasibility study of a low-fat diet. *Arch Intern Med.* 1990; 150:421–427.
24. Block G, Woods M, Potoshy A, Clifford C. Validation of a self-administered diet history questionnaire using multiple diet records. *J Clin Epidemiol.* 1990;43:1327–1335.
25. National Cancer Institute, Division of Cancer Prevention and Control. *Health Habits and History Questionnaire: Diet History and Other Risk Factors* [Personal computer system packet]. Bethesda, Md: National Cancer Institute; 1988.
26. Miller DC. *Handbook of Research Design and Social Measurement.* 3rd ed. New York, NY: Longman; 1977:230–242.
27. Armitage P, Berry G. *Statistical Methods in Medical Research.* 2nd ed. Oxford, England: Blackwell Scientific Publications; 1987.
28. Kleinbaum DG, Kupper LL, Muller KE. *Applied Regression Analysis and Other Multivariable Methods.* Boston, Mass: PWS-KENT Publishing Co; 1988.
29. Block G, Rosenberger WF, Patterson BH. Calories, fat, and cholesterol: intake patterns in the US population by race, sex, and age. *Am J Public Health.* 1988;78:1150–1155.
30. Stephen AM, Wald NJ. Trends in individual consumption of dietary fat in the United States, 1920–1984. *Am J Clin Nutr.* 1990; 52:456–469.
31. Gorder DD, Dolecek TA, Coleman GG, et al. Dietary intake in the Multiple Risk Factor Intervention Trial (MRFIT): nutrient and food group changes over 6 years. *J Am Diet Assoc.* 1986;86:744–751.
32. Burr ML, Fehily AM, Gilbert JF, et al. Effects of changes in fat, fish and fibre intakes on death and myocardial reinfarction: diet and reinfarction trial (DART). *Lancet.* 1989;11:757–761.
33. Gorbach SL, Morrill-LaBrode A, Woods MN, et al. Changes in food patterns during a low-fat dietary intervention in women. *J Am Diet Assoc.* 1990;90:802–809.
34. Schafer RB. Factors affecting food behavior and the quality of husbands' and wives' diets. *J Am Diet Assoc.* 1978;72:138–143.

Weight Loss Attempts in Adults: Goals, Duration, and Rate of Weight Loss

David F. Williamson, PhD, Mary K. Serdula, MD, Robert F. Anda, MD, Alan Levy, PhD, and Tim Byers, MD

ABSTRACT

Objectives. Although attempted weight loss is common, little is known about the goals and durations of weight loss attempts and the rates of achieved weight loss in the general population.

Methods. Data were collected by telephone in 1989 from adults aged 18 years and older in 39 states and the District of Columbia. Analyses were carried out separately for the 6758 men and 14 915 women who reported currently trying to lose weight.

Results. Approximately 25% of the men respondents and 40% of the women respondents reported that they were currently trying to lose weight. Among men, a higher percentage of Hispanics (31%) than of Whites (25%) or Blacks (23%) reported trying to lose weight. Among women, however, there were no ethnic differences in prevalence. The average man wanted to lose 30 pounds and to weigh 178 pounds; the average woman wanted to lose 31 pounds and to weigh 133 pounds. Black women wanted to lose an average of 8 pounds more than did White women, but Black women's goal weight was 10 pounds heavier. The average rate of achieved weight loss was 1.4 pounds per week for men and 1.1 pounds per week for women; these averages, however, may reflect only the experience of those most successful at losing weight.

Conclusions. Attempted weight loss is a common behavior, regardless of age, gender, or ethnicity, and weight loss goals are substantial; however, obesity remains a major public health problem in the United States. (*Am J Public Health.* 1992; 82:1251–1257)

Introduction

Voluntary weight loss is a common practice in the United States and appears to be increasing in prevalence. In national surveys conducted between 1950 and 1966, about 7% of men and 14% of women reported that they were trying to lose weight.[1,2] A 1978 survey found that 16% of adults were trying to lose weight.[1,3] More recent data from the 1985 National Health Interview Survey suggest that the prevalence of attempted weight loss has increased substantially; about 25% of men and 45% of women reported that they were trying to lose weight.[4]

Although attempts to lose weight are common, data on specific aspects of this behavior in the general population are limited.[1] Much of our knowledge about voluntary weight loss comes from clinical studies of self-selected subjects, many of whom are severely obese. The findings from such studies may produce biased estimates if applied to the general population.[1,5,6] The purpose of this study is to describe the age- and race-specific prevalence of weight loss attempts, the goals and duration of such attempts, and the rate of weight loss achieved in a population-based, multistate health practices survey.

Methods

Data Source

Data for this analysis are from the 1989 Behavioral Risk Factor Surveillance System (BRFSS) cross-sectional survey of 39 states and the District of Columbia. The BRFSS is a joint effort between the Centers for Disease Control (CDC) and participating state health departments. The BRFSS is a random-digit-dialed telephone survey of each state's noninstitutionalized civilian population aged 18 years and older with telephones. Each participating state conducts approximately 100 interviews per month throughout the year, using a multistage cluster design.[7] All states use an identical core questionnaire, and CDC staff provide standardized training in survey operations. In 1989 the median state response rate (the ratio of completed interviews to the sum of completed interviews and refusals) was 82%. A detailed technical description of the BRFSS has been presented elsewhere.[8]

The BRFSS telephone interview included 65 core questions and took about 25 minutes to administer. The questions on weight control practices were asked of all respondents near the middle of the interview, and followed questions on seat belt use, hypertension, and physical activity. Questions on current self-reported weight and height were asked at the end of the interview.

Weight Control Status

Respondents were first asked, "Are you now trying to lose weight?" If they answered no, they were asked, "Are you now trying to maintain your current

David F. Williamson, Mary K. Serdula, Robert F. Anda, and Tim Byers are with the National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control, Atlanta, Ga. Alan Levy is with the Division of Consumer Studies, Food and Drug Administration, Washington, DC.

Requests for reprints should be sent to David F. Williamson, PhD, Division of Nutrition (K-26), Centers for Disease Control, 1600 Clifton Rd, NE, Atlanta, GA 30333.

This paper was submitted to the Journal October 8, 1991, and accepted with revisions April 7, 1992.

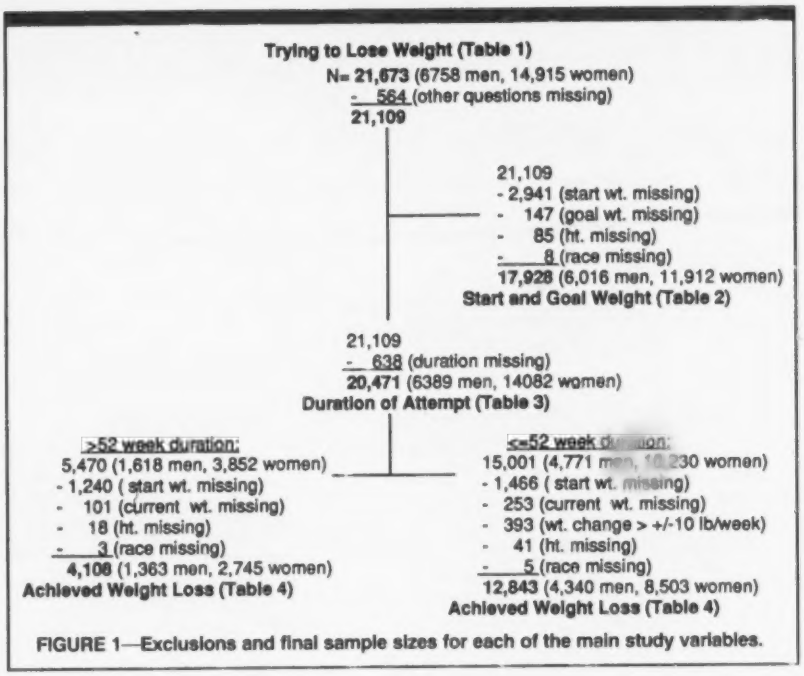


FIGURE 1—Exclusions and final sample sizes for each of the main study variables.

weight, that is, to keep from gaining weight?" Persons who responded no to this question were considered to be doing nothing about their weight. Of the 66 867 respondents who completed the telephone interview, 153 (0.2%) did not answer yes or no to the question of whether they were trying to lose weight, 1680 (2.5%) reported that they were trying to gain weight, and 397 (0.6%) did not answer the question as to whether they were trying to maintain their weight. This left 64 637 persons who could be classified into one of the three groups according to their current weight control status: trying to lose weight (n = 21 673; 33.5%), trying to maintain weight (n = 18 253; 28.2%), or doing neither (24 711; 38.3%).

Duration of Weight Loss Attempt

Those who reported that they were trying to lose weight were asked, "About how long ago did you begin your current attempt to lose weight?" We expressed duration of attempted weight loss in terms of weeks; in some analyses persons who reported trying to lose weight for more than 52 weeks were combined with those who reported that they were "always trying to lose weight" and were analyzed as a separate category of duration.

Weight Loss Goal

Persons trying to lose weight were next asked, "About how much did you weigh when you began your current attempt to lose weight?" and "How much would you like to weigh?" These questions were used to estimate the respondents' mean weight at the start of their weight loss attempt and their mean goal weight. We computed the difference between the starting weight and the goal weight as the weight loss goal and expressed both in pounds and in relative terms by dividing the weight loss goal by the starting weight to obtain a percentage weight loss goal. For some analyses we categorized respondents according to four categories of body mass index (BMI: kg/m^2): < 26, 26 to < 30, 30 to < 35, ≥ 35. We classified persons with a BMI ≥ 30 as obese.[9]

Achieved Weight Loss

We computed the difference between the starting weight and the current weight as the achieved weight "loss" (this included some respondents who had a net gain in weight). The achieved weight loss was computed separately for durations > 52 weeks (including those who reported "always trying to lose weight") and for durations of ≤ 52 weeks. For the respondents who reported a duration of ≤ 52 weeks, we also estimated the rate of achieved weight loss, expressed as pounds per week. We considered a weight change of more than ±10 pounds per week to be biologically implausible, and the records of persons who reported such high rates were excluded. We also estimated the prevalence of achieved weight loss of more than 2 pounds per week, a rate that is considered potentially harmful.[10,11]

Final sample sizes. This analysis focuses primarily on the 21 673 respondents who reported that they were currently trying to lose weight. One state did not ask any of the questions on duration, starting weight, or goal weight; respondents in that state (n = 564, 2.6% of the total) were excluded from further analyses, leaving 21 109 respondents to answer the remaining questions. Subsequent exclusions and final sample sizes for each of the main study variables are shown in Figure 1.

Statistical analysis. We report all results separately for men and women. Because of the large sample sizes, we have not emphasized statistical testing; however, *P* value estimates are given in the accompanying tables. In some analyses we computed age- and race-specific prevalence estimates and their 95% confidence limits (CL).[12] We used the Pearson chi square to test for homogeneity of the prevalence estimates across age and race strata. Weighted least squares estimation was used to statistically test for interaction between age and race on the prevalence estimates.[13] For those who reported trying to lose weight for 52 weeks or less, we computed the age- and race-specific median duration (in weeks) of weight loss and the 95% CL.[14] Age- and race-specific mean starting and goal weights and mean weight loss goal with 95% CL were estimated by analysis of variance,[13] which also adjusted for height. Race-specific estimates were also adjusted for age, as were estimates of achieved weight loss and rate of weight loss. Although a higher proportion of Whites (55%) than of Blacks (41%) and Hispanics (44%) had education beyond high school, further adjustment for education had little effect on race-specific estimates.

Results

The mean age of men respondents was 44 years; 85% were White, 63% were married, and 53% had more than 12 years of education. Their mean BMI was 25.7, which is comparable to the most recent national estimate of mean BMI in men of 25.3.[15] The mean age of women respondents was 46 years; 84% were White, 53% were married, and 47% had more than 12 years of education. Their mean current BMI was 24.2, somewhat lower than the national estimate of 25.0.[15]

Weight Control Status

About 25% of the men (95% CI = 24%, 26%) reported that they were currently trying to lose weight (Table 1). The peak prevalence of trying to lose weight occurred among those aged 40 to 49 years (30%) and among Hispanics (31%). BMI, however, had a much stronger relationship to the prevalence of trying to lose weight than did age or ethnicity. The prevalence increased more than fivefold between men with a BMI of < 26 (13%) and those with a BMI of > 35 (65%).

Among women, 39% (95% CI = 38%, 40%) reported that they were currently trying to lose weight (Table 1). The association with age was similar to that seen in men; in women, however, ethnicity was unrelated to the prevalence of trying to lose weight. For women as for men, BMI had the strongest relationship to the prevalence of trying to lose weight.

TABLE 1—Weight Control Status Prevalence (Expressed as Percentages) by Selected Characteristics of Respondents to the 1989 Behavioral Risk Factor Surveillance Survey

	Trying to Lose, %	Trying to Maintain, %	Neither, %
	Men		
	(n = 6758)	(n = 7662)	(n = 12 318)
Age,[a] y			
18–29	20.2	28.4	51.4
30–39	28.4	29.8	43.8
40–49	30.4	29.3	40.4
50–59	29.9	28.0	42.1
60–69	28.0	27.8	44.2
70+	16.0	27.0	57.0
Race[b]			
White	25.1	28.5	46.4
Black	23.4	27.0	49.7
Hispanic[c]	30.8	29.0	40.2
Other	25.6	35.0	39.2
Current BMI,[d] kg/m^2			
<26	12.6	31.6	55.8
26–29.9	40.3	27.1	32.7
30–34.9	56.1	18.7	25.2
≥35	65.1	15.1	19.7
Total	25.2	28.7	46.1
	Women		
	(n = 14 915)	(n = 7662)	(n = 12 313)
Age,[e] y			
18–29	40.1	28.8	31.1
30–39	44.0	27.9	28.1
40–49	46.9	27.2	25.9
50–59	43.9	27.8	28.4
60–69	37.3	28.4	34.2
70+	21.4	27.3	51.4
Race[f]			
White	39.1	28.1	32.8
Black	41.1	25.3	33.6
Hispanic[c]	39.7	29.0	31.3
Other	40.6	28.9	30.5
Current BMI,[g] kg/m^2			
<26	29.2	33.0	37.8
26–29.9	58.5	19.1	22.4
30–34.9	63.9	17.4	18.7
≥35	68.6	12.8	18.6
Total	39.3	27.9	32.3

Note. Each percentage is a percentage of its row. Percentages may not add to 100 because of rounding. The chi-square statistics apply only to persons trying to lose weight and are a test for homogeneity of the prevalences across levels of each characteristic. BMI = body mass index.

[a] χ^2 (df, 5) = 319.6, $P < .001$.
[b] χ^2 (df, 3) = 24.4, $P < .001$.
[c] Of those reporting Hispanic origin, 51.0% were from the following states: Arizona (5.7%), California (14.0%), New Mexico (11.0%), Texas (9.5%), Florida (5.8%), and New York (5.0%).
[d] χ^2 (df, 3) = 3942.4, $P < .001$.
[e] χ^2 (df, 5) = 1039.2, $P < .001$.
[f] χ^2 (df, 3) = 6.1, $P < .11$.
[g] χ^2 (df, 3) = 3157.7, $P < .001$

Weight Loss Goal

Among men trying to lose weight, the mean weight at the start of the weight loss attempt was 209 pounds; the mean goal weight was 178 pounds, for a mean weight loss goal of 30 pounds (13.4% of the starting body weight) (Table 2). The youngest age group (18 to 29 years) had lower mean starting and goal weights than did the older age groups. Compared with Whites, Black men had higher mean starting and goal weights, but they also had a slightly higher mean goal weight loss, as did Hispanic men.

Among women trying to lose weight, the mean starting weight was 164 pounds; the mean goal weight was 133 pounds, for a mean weight loss goal of 32 pounds (17.6% of the starting weight) (Table 2). As was true for men, the youngest group had lower mean starting and goal weights than did the older groups. Among women, there were substantial differences in weights and weight loss goals between Whites and Blacks and, to a lesser degree, between Whites and Hispanics. Black women had a 19-pound-heavier mean starting weight and an 11-pound-heavier mean goal weight, but an 8-pound-greater mean weight loss goal, than did their White counterparts.

We found a strong association between gender and starting weight and goal weight (expressed as BMI) (data not shown). Among those in the starting BMI category of 30 to 34.9, 39% of men, compared with 80% of women, had a weight loss goal that would place them in the leanest goal BMI category (< 26). Among those in the heaviest starting BMI category (≥ 35), 19% of men, compared with 59% of women, wanted to end up in the leanest category. Thus, obese women are two to three times more likely than obese men to have a weight loss goal that would place them in the leanest body weight category.

Duration of Weight Loss Attempts

Among men trying to lose weight, the prevalence of long-term weight loss attempts (defined as trying to lose weight for more than a year or "always trying to lose weight") was 25% and was directly related to age (Table 3). There was a mod-

TABLE 2—Mean Starting Weight, Goal Weight, and Weight Loss Goal (In Pounds) of Persons Currently Trying to Lose Weight

		Starting Weight		Goal Weight		Goal Loss		
	No.	Mean	95% CI	Mean	95% CI	Mean	95% CI	%[a]
Men								
Age, y								
18–29	1124	202.4	± 1.8	174.8	± 1.0	27.7	± 1.4	12.8
30–39	1633	206.9	± 1.6[b]	177.5	± 0.8[b]	29.5	± 1.2[b]	13.4
40–49	1301	210.9	± 1.8[b]	179.4	± 0.8[b]	31.5	± 1.2[b]	14.0
50–59	843	212.9	± 2.2[b]	180.5	± 1.2[b]	32.4	± 1.6[b]	14.3
60–69	771	211.2	± 2.2[b]	180.4	± 1.2[b]	30.8	± 1.6[b]	13.7
70+	344	209.0	± 3.3[b]	179.9	± 1.8[b]	29.1	± 2.4	13.2
Race								
White	5059	207.7	± 0.8	178.0	± 0.4	29.7	± 0.6	13.4
Black	402	216.1	± 3.0[b]	182.8	± 1.6[b]	33.2	± 2.2[b]	14.7
Hispanic	339	212.2	± 3.3[b]	178.7	± 1.8	33.4	± 2.4[b]	14.5
Other	216	205.9	± 4.1	176.6	± 2.2	29.3	± 3.1	12.9
All	6016	208.5	± 1.0	178.3	± 0.6	30.1	± 0.6	13.4
Women								
Age, y								
18–29	2753	154.5	± 1.2	125.3	± 0.6	29.2	± 1.0	17.3
30–39	3156	162.4	± 1.2[b]	129.6	± 0.4[b]	32.8	± 1.0[b]	18.3
40–49	2169	166.1	± 1.4[b]	132.8	± 0.6[b]	33.3	± 1.2[b]	18.1
50–59	1513	170.7	± 1.6[b]	137.1	± 0.8[b]	33.6	± 1.4[b]	17.9
60–69	1448	171.9	± 1.8[b]	141.1	± 0.8[b]	30.8	± 1.4	16.5
70+	873	169.5	± 2.2[b]	142.6	± 1.0[b]	27.0	± 1.8	14.9
Race								
White	9836	162.0	± 0.8	131.3	± 0.2	30.7	± 0.6	17.3
Black	1182	181.1	± 1.8[b]	142.5	± 0.8[b]	38.6	± 1.4[b]	19.7
Hispanic	520	167.0	± 2.8[b]	134.2	± 1.2	32.8	± 2.2	18.1
Other	374	159.3	± 3.3	130.7	± 1.4	28.6	± 2.5	16.1
All	11 912	164.0	± 0.6	132.5	± 0.4	31.5	± 0.4	17.6

Note. All estimates are adjusted for height. Race-specific estimates are also adjusted for age.
CI = confidence interval.
[a]Percentage of starting body weight.
[b]Statistically different ($P < .05$) from 18–29 age group and from Whites.

est relationship between race and long-term weight loss attempts; prevalence ranged from 23% among Black men to 31% among men of "other" race. If the prevalence of long-term weight loss attempts is applied to the total sample of men, then 6% (0.253 × 0.252 × 100) of all men were trying to lose weight for more than a year.

Among women trying to lose weight, the prevalence of long-term weight loss attempts was 27%, which was similar to that for men (Table 3). There was also a consistent increase in prevalence with increasing age. In contrast to men, however, for women there was a strong association between race and the prevalence of long-term weight loss attempts, which ranged from 18% among Blacks to 29% among Whites. In terms of the total sample, 11% of all women were trying to lose weight for more than a year.

We also estimated the median duration of the current weight loss attempt for those trying to lose weight for a year or less (data not shown). Among men, the median duration of the current weight loss attempt was 6 weeks (95% CL = ± 0.3 weeks). The median duration increased with increasing age, from 4 weeks among those in the 18 to 29 age group to 12 weeks among men in the 70+ age group. The median duration was 2 weeks shorter for White men than for men of other races. Among women, the median duration of the current weight loss attempt was 4 weeks (95% CL = ± 0.2 weeks). The median duration was 4 weeks for women younger than 50 years of age and rose to 8 weeks for women aged 50 and older. Women of "other" race had a median duration of 8 weeks, compared with 4 weeks for the three racial groups.

Achieved Weight Loss

Among men trying to lose weight for more than 1 year, the mean achieved weight loss was 12 pounds (Table 4). There was substantial variation in weight loss across age categories, with the highest weight loss (18 pounds) occurring in those aged 18 to 29 years and the smallest loss (10 pounds) occurring in those aged 30 to 39 years. There was a strong relationship between race and achieved weight loss; Hispanic men reported the largest weight loss (19 pounds) and men classified as "other" race reported the smallest loss (11 pounds).

Among men trying to lose weight for 1 year or less, the mean achieved weight loss was 10 pounds and varied strongly with age, ranging from 9 pounds for men aged 30 to 39 to 12 pounds for men aged 60 to 69 (Table 5). The mean rate of weight loss was 1.4 pounds per week; this figure also varied substantially with age, ranging from a loss of 1.6 pounds per week for men aged 18 to 29 to 0.9 pounds for men aged 70 and older. There was little relationship between race and the rate of weight loss in men.

Among women who reported trying to lose weight for more than a year, the mean achieved weight loss was 8.8 pounds (Table 4). As was seen with men, the largest loss (12 pounds) occurred in women aged 18 to 29 and the smallest loss (8 pounds) was among women aged 40 to 49. There was little difference by race in achieved weight loss for women.

Among women trying to lose weight for a year or less, the mean achieved weight loss was 8 pounds and was unrelated to age (Table 5). However, the achieved weight loss was nearly 1.5 pounds lower for Black women (7.2 pounds) than for White women (8.6 pounds). The mean rate of weight loss for women was 1.1 pounds per week and was inversely related to age, ranging from a loss of 1.1 pounds per week for women aged 18 to 29 to a loss of 0.7 pounds per week for women aged 70 years and older. Race was unrelated to the rate of weight loss in women.

The prevalence of a rate of achieved weight loss in excess of 2 pounds per week was 25% for men and 18% for women (data not shown). The prevalence was consistently 5 to 10 percentage points higher for men than for women in all age groups and was inversely related to age. The prevalence ranged from a high of 31% for men aged 18 to 29 years to a low of 10% for women aged 70 years and older. The prevalence was not related to race for either men or women.

Discussion

Using data from a geographically and ethnically diverse sample of adults in the general population, we found that approximately 25% of men and 40% of women were trying to lose weight, with the peak prevalence occurring among persons in their 40s. These findings are similar to those from a recent nationally representative in-person survey of health practices, in which 25% of men and 44% of women reported that they were currently trying to lose weight.[4]

Approximately 25% of both men and women who were currently trying to lose weight could be defined as long-term dieters. Thus, in the total sample, the prevalence of long-term dieting is about 6% of all men and 11% of all women. In a national 1985 Gallup Poll,[16] 16% of women considered themselves "perpetual dieters." This higher estimate may have been due to different definitions of long-term dieting. We found no published estimates of the prevalence of long-term weight loss attempts by men.

The average person trying to lose weight wanted to lose about 30 pounds. This figure is similar to that found in earlier studies of self-selected samples. Ureda[17] found that the median weight loss goal of men and women volunteers in a university-based weight loss program was 35 pounds, and Muhlenkamp and Nelson[18] reported an average weight loss goal of 33 pounds in a sample of clients of a commercial weight loss program (gender not given).

In absolute terms, the weight loss goal was very similar for both men and women, but relative to their starting weights, the goal of women was about one-third higher than that of men (17% vs 13% of body weight). The apparent desire of women to lose a larger proportion of their weight was also reflected in the relationship between starting BMI and goal BMI: women in the heaviest BMI categories at the start of their weight loss attempt were two to three times more likely than men to want to end up in the leanest BMI category. We found no published studies that directly compared weight loss goals of men and women. However, our findings support the view that women are under greater pressure than men to lose weight.[19]

The achieved weight loss in this study was similar for long-term dieters and shorter-term dieters. Regardless of duration, the average man had lost 10 to 12 pounds and the average women had lost 8 to 9 pounds. It may be that long-term dieting is simply a marker for difficulty in losing weight. It is also possible that long-term dieters are more likely to lose and regain some of their weight cyclically during any one attempt. However, because our study had information only on starting and current weight, the true underlying variability in weights of long-term dieters may have been masked.

The higher amount and higher rate of achieved weight loss reported by men are consistent with the results of clinical studies[20,21] and are a function of the higher average starting weights of men.[22] After

TABLE 3—Prevalence of Weight Loss Attempt of More than 1 Year's Duration for Men and Women, by Age and Race

		Prevalence of Current Attempt > 1 Year Duration					
Age,[a] y	No.[b]	%[c]	95% CI	Race[d]	No.[b]	%[c]	95% CI
			Men				
18–29	1183	12.6	± 1.9	White	5361	25.3	± 0.5
30–39	1732	19.1	± 1.9	Black	430	22.8	± 2.0
40–49	1383	26.8	± 2.3	Hispanic	361	25.2	± 2.3
50–59	892	31.2	± 3.0	Other	232	30.6	± 3.0
60–69	828	38.0	± 3.3				
70+	371	47.2	± 5.1				
All	6389	25.3	± 1.1				
			Women				
18–29	3116	15.4	± 1.3	White	11 710	28.8	± 0.4
30–39	3734	23.1	± 1.4	Black	1331	18.3	± 1.1
40–49	2638	27.2	± 1.7	Hispanic	607	22.7	± 1.7
50–59	1803	34.1	± 2.2	Other	430	23.5	± 2.0
60–69	1699	40.1	± 2.3				
70+	1092	45.2	± 3.0				
All	14 082	27.3	± 0.7				

Note. CI = confidence interval.
[a] χ^2 (*df*, 5) = 318.8, $P < .0001$, for men; χ^2 (*df*, 5) = 612.2, $P < .0001$, for women.
[b] The age- and race-specific sample sizes represent the number of respondents currently trying to lose weight who reported the duration of their weight loss attempt.
[c] The % column represents the proportion of those respondents whose attempt lasted longer than 1 year.
[d] χ^2 (*df*, 3) = 4.9, $P = .18$, for men; χ^2 (*df*, 3) = 75.8, $P < .0001$, for women.

TABLE 4—Achieved Weight Loss (in Pounds) Among Persons Trying to Lose Weight for More than 1 Year

Age, y	No.	Mean	95% CI	Race	No.	Mean	95% CI
			Men				
18–29	120	−18.0	± 3.9	White	1140	−11.5	± 1.2
30–39	285	−9.9	± 2.5*	Black	79	−14.9	± 4.7
40–49	310	−11.5	± 2.4*	Hispanic	80	−18.8	± 4.7*
50–59	236	−11.1	± 2.7*	Other	64	−10.5	± 5.3
60–69	263	−12.8	± 2.5*				
70+	149	−12.9	± 3.5				
All	1363	−12.1	± 1.2				
			Women				
18–29	368	−11.8	± 2.2	White	2398	−8.9	± 0.8
30–39	630	−8.8	± 1.8*	Black	170	−7.8	± 3.3
40–49	483	−7.6	± 2.0*	Hispanic	102	−7.2	± 4.3*
50–59	430	−8.3	± 2.0*	Other	75	−9.7	± 4.9
60–69	492	−8.8	± 2.0*				
70+	342	−8.0	± 2.4*				
All	2745	−8.8	± 0.8				

Note. Estimates for all ages and age-specific estimates are adjusted for height only. CI = confidence interval.
* Statistically different ($P < .05$) from 18–29 age group and from Whites.

TABLE 5—Achieved Weight Loss (in Pounds) Among Persons Trying to Lose Weight for 1 Year or Less

		Total Loss		Loss per Week				Total Loss		Loss per Week	
Age, y	No.	Mean	95% CI	Mean	95% CI	Race	No.	Mean	95% CI	Mean	95% CI
						Men					
18–29	932	−9.4	± 0.8	−1.6	± 0.1*	White	3676	−9.8	± 0.4	−1.4	± 0.1
30–39	1258	−8.9	± 0.7	−1.3	± 0.1*	Black	281	−9.2	± 1.4	−1.2	± 0.2
40–49	940	−9.9	± 0.8	−1.4	± 0.1*	Hispanic	237	−11.8	± 1.6*	−1.3	± 0.2
50–59	557	−11.2	± 1.1*	−1.2	± 0.1*	Other	146	−11.4	± 2.2	−1.4	± 0.3
60–69	479	−12.1	± 1.2*	−1.4	± 0.2*						
70+	174	−10.3	± 2.0	−0.9	± 0.3*						
All	4340	−9.9	± 0.4	−1.4	± 0.1						
						Women					
18–29	2219	−8.3	± 0.5	−1.2	± 0.1	White	6939	−8.6	± 0.4	−1.1	± 0.1
30–39	2382	−8.4	± 0.5	−1.1	± 0.8*	Black	897	−7.2	± 0.9*	−1.1	± 0.1
40–49	1574	−8.7	± 0.6	−1.1	± 0.1*	Hispanic	388	−7.9	± 1.3	−1.2	± 0.2
50–59	991	−8.4	± 0.8	−1.1	± 0.1*	Other	279	−7.4	± 1.6	−0.9	± 0.2
60–69	854	−7.9	± 0.8	−0.9	± 0.1*						
70+	483	−8.0	± 1.2	−0.7	± 0.1*						
All	8503	−8.4	± 0.3	−1.1	± 0.03						

Note. Estimates for all ages and age-specific estimates are adjusted for height only. CI = confidence interval.
*Statistically different ($P < .05$) from 18–29 age group and from Whites.

adjusting for starting BMI, we found that the rate of weight loss for men and women was identical at 1.3 pounds per week.

Ethnicity had a modest association with the prevalence of attempted weight loss for men, with Hispanic men reporting the highest prevalence. Black and Hispanic men also tended to weigh more than White men when they started their attempt to lose weight. After controlling for duration, we found that the rate of weight loss was nearly the same for all ethnic groups of men. Previous studies suggest either that there is no difference in "dieting behavior"[23] between Hispanic and Anglo men or that Hispanic men are less inclined to diet to lose weight.[24] Our finding of a higher prevalence of attempted weight loss by Hispanic men may reflect the higher socioeconomic status of Hispanics from telephone-owning households.

Among women, however, we found no association between the prevalence of attempted weight loss and ethnicity. A lower prevalence of attempted weight loss has been proposed as a potential cause of the higher levels of obesity in Black women.[25] Although studies done in the 1960s suggested that Black women were less likely than White women to try to lose weight,[25] more recent studies indicate that there is less difference between Black and White women in the prevalence of attempting to lose weight.[25–28] We did find differences between White and Black women in weight loss goals. Although Black women had higher starting and goal body weights than did White women, Black women still wanted to lose about 8 pounds more than did their White counterparts. We also found ethnic differences in women in the prevalence of long-term attempted weight loss; White women were nearly 60% more likely than Black women to report that they had been trying to lose weight for a year or more. Black women reported lower achieved weight losses than did White women, which was also observed in earlier studies.[29,30] After we adjusted for duration, however, the rate of weight loss for Black and White women was identical. Taken together, our findings do not suggest that Black or Hispanic women are markedly different from White women in their weight loss behavior.

Although 39 states and the District of Columbia were included in the sample, this survey was not designed to represent the US population. The sample was also limited to telephone-owning households; although only about 4% to 7% of US households do not own telephones, weight control practices may differ depending on telephone ownership. This sampling bias may be particularly important for ethnic minorities, for whom telephone ownership is lower than in the general population.[31] We know of no study that has examined differences in weight control practices between persons from households with and without telephones. Previous studies have generally reported that there are only modest differences in other health practices between persons from households with telephones and those without telephones.[31,32]

This study had a survey nonresponse rate of 18%. In a study of nonrespondents to a general-topic telephone survey, however, it was found that the survey's population estimates were only slightly altered by including data collected from the initial nonrespondents.[33] In addition, a recent study of the BRFSS found no statistically significant differences in the mean BMI or in the prevalence of current weight loss attempts between initial responders and nonresponders.[34]

A related issue is that of item nonresponse, which was highest for the question on weight at the start of the weight loss attempt (see Figure 1). It may be that those who had very high weights were less willing to answer this question because of embarrassment, or because they were unsuccessful at losing weight. If this was true, then our estimates of mean starting weights may be too low. Differential item nonresponse may also account for the similarity in achieved weight loss between long-term and short-term dieters. For long-term dieters, nearly 23% of starting weights were missing, compared with 10% for short-term dieters. Hence, our estimates of achieved weight loss must be viewed cautiously.

In this study weight and height were self-reported. Self-reports produce downwardly biased estimates of overweight,[35] and this bias may be greater in telephone surveys.[36] Hence, some persons were probably misclassified in lower BMI categories. Past body weight is generally reported with reasonable accuracy, even

when the recall period is as long as 40 years.[37,38] We believe, however, that persons who are attempting to lose weight are more likely than the general population to know their current and previous weights.

Our study suggests that nearly one of every three adults in the United States is currently trying to lose weight, making voluntary weight loss one of the most common health-related practices in the United States today. Although more than $30 billion is spent annually on weight loss,[39] the long-term health effects of weight loss remain unclear.[40,41] Thus, it is important to provide better information to the public health community about this poorly understood but highly prevalent practice. □

Acknowledgments

We would like to thank Lee Sanderson, PhD, and Emma Frazier, PhD, of the Office of Surveillance and Analysis, CDC, for their help in using these data. In addition, we acknowledge the state BRFSS coordinators, without whose cooperation this work would not have been possible.

Portions of this paper were presented at the 26th annual meeting of the US Public Health Service Professional Association, Atlanta, Ga, May 26–29, 1991.

References

1. Jeffery RW, Folsom AR, Russel VL, et al. Prevalence of overweight and weight loss behavior in a metropolitan adult population: the Minnesota heart survey experience. *Am J Public Health.* 1984;74:349–352.
2. Dwyer JT, Mayer J. Potential dieters: who are they? *J Am Diet Assoc.* 1970;56:510–514.
3. *Health Maintenance: Weight and Obesity.* Newport Beach, Calif: Pacific Mutual Life Insurance Co; 1978.
4. Stephenson MG, Levy AS, Sass NL, McGarvey WE. 1985 NHIS findings: nutrition knowledge and baseline data for the weight-loss objectives. *Public Health Rep.* 1987;102:61–67.
5. Blackburn GL, Kanders BS. Medical evaluation and treatment of the obese patient with cardiovascular disease. *Am J Cardiol.* 1987;60:55G–58G.
6. Drewnowski A. Human obesities and the response to diet. Presented at the Harvard CMI course on treatment of obesity: a multidisciplinary approach; October 21, 1989; Boston, Mass.
7. Waksberg JS. Methods for random digit dialing. *J Am Stat Assoc.* 1978;73:40–46.
8. Remington PL, Smith MY, Williamson DF, Anda RF, Gentry EM, Hogelin GC. Design, characteristics, and usefulness of state-based Behavioral Risk Factor Surveillance: 1981–1987. *Public Health Rep.* 1988;103:366–375.
9. Committee on Diet and Health, National Research Council. *Diet and Health: Implications for Reducing Chronic Disease Risk.* Washington, DC: National Academy Press; 1989.
10. Michigan Health Council (Petersmarck KA, ed.) Toward safe weight loss: Recommendations for adult weight loss programs in Michigan. Lansing, Mich: The Center for Health Promotion, Michigan Department of Public Health; December 1990.
11. Council on Scientific Affairs. Treatment of obesity in adults. *JAMA.* 1988;260:2547–2551.
12. Fleiss JL. *Statistical Methods for Rates and Proportions.* New York, NY: John Wiley & Sons, Inc; 1981:13–14.
13. SAS Institute Inc. *SAS User's Guide: Statistics, Version 5 Edition.* Cary, NC: SAS Institute Inc; 1985. Proc Freq: 403–432, Proc Catmod: 171–254, Proc GLM: 433–506.
14. Velleman PF, Hoaglin DC. Applications, basics, and computing of exploratory data analysis. Boston, Mass: Duxbury Press; 1981:79–81.
15. Najjar MF, Rowland M. *Anthropometric Reference Data and Prevalence of Overweight, United States, 1976–80.* Washington, DC: US Government Printing Office; October 1987. Vital and Health Statistics. Series 11, No. 238. US Depart of Health and Human Services publication PHS 87-1688.
16. Toufexis A. Dieting: the losing game. *Time.* January 20, 1986;127:54–60.
17. Ureda JR. The effect of contract witnessing on motivation and weight loss in a weight control program. *Health Educ Q.* 1980;7:163–185.
18. Mulhlenkamp AF, Nelson AM. Health locus of control, values, and weight reduction behavior. *J Psychosoc Nurs Ment Health Serv.* 1981;19:21–25.
19. Rodin J, Silberstein L, Striegel-Moore R. Women and weight: a normative discontent. In: Sonderegger KB, ed. *Psychology and Gender. Nebraska Symposium on Motivation.* Vol. 32. Lincoln, Neb: University of Nebraska Press; 1985:267–707.
20. Kramer FM, Jeffery RW, Forster JL, Snell MK. Long-term follow-up of behavioral treatment for obesity: patterns of weight regain among men and women. *Int J Obesity.* 1989;13:123–136.
21. Brownell KD, Steen SN. Modern methods for weight control: the physiology and psychology of dieting. *Physician Sportsmed.* 1987;15:31–40.
22. Van Itallie TB. Conservative approaches to treatment. In: Bray GA, ed. *Obesity in America.* Washington, DC: US Government Printing Office; 1979:164–178. US Dept of Health, Education, and Welfare publication NIH 79-359.
23. Stern MP, Pugh JA, Gaskill SP, Hazuda HP. Knowledge, attitudes, and behavior related to obesity and dieting in Mexican Americans and Anglos: the San Antonio Heart Study. *Am J Epidemiol.* 1982;115:917–928.
24. Haffner SM, Stern MP, Mitchell BD, Hazuda HP. Predictors of obesity in Mexican Americans. *Am J Clin Nutr.* 1991;53:1571S–1576S.
25. Kumanyika S. Obesity in black women. *Epidemiol Rev.* 1987;9:31–50.
26. Dawson DA. Ethnic differences in female overweight: data from the 1985 National Health Interview Survey. *Am J Public Health.* 1988;78:1326–1329.
27. Gillum RF. Overweight and obesity in black women: a review of published data from the National Center for Health Statistics. *J Natl Med Assoc.* 1987;79:865–871.
28. Wing RR, Kuller LH, Bunker C, et al. Obesity, obesity-related behaviors and coronary heart disease risk factors in black and white premenopausal women. *Int J Obesity.* 1989;13:511–519.
29. Kumanyika SK, Obarzanek E, Stevens VJ, Hebert PR, Whelton PK. Weight-loss experience of black and white participants in NHLBI-sponsored clinical trials. *Am J Clin Nutr.* 1991;53:1631S–1638S.
30. Kahn HS, Williamson DF. Is race associated with weight change in US adults after adjustment for income, education, and marital factors? *Am J Clin Nutr.* 1991;53:1566S–1570S.
31. Marcus AC, Crane LA. *The Validity and Value of Health Survey Research by Telephone.* New York, NY: The Commonwealth Fund; 1986.
32. Thornberry O. *An Experimental Comparison of Telephone and Personal Health Interview Surveys.* Washington, DC: US Government Printing Office; August 1987. Vital and Health Statistics. Series 2, No. 106. US Dept of Health and Human Services publication PHS 87-1380.
33. Marcus AC, Crane LA. Telephone surveys in public health research. *Med Care.* 1986;24:97–112.
34. Shea S, Stein AD, Lantigua R, Basch CE. Reliability of the behavioral risk factor survey in a triethnic population. *Am J Epidemiol.* 1991;133:489–500.
35. Nieto-Garcia FJ, Bush TL, Keyl PM. Body mass definitions of obesity: sensitivity and specificity using self-reported weight and height. *Epidemiol.* 1990;1:146–152.
36. Jackson C, Jatulis DE, Fortmann SP. The Behavioral Risk Factor Survey and the Stanford Five-City Project Survey: a comparison of cardiovascular risk behavior estimates. *Am J Public Health.* 1992;82:412–416.
37. Stevens J, Keil JE, Waid LR, Gazes PC. Accuracy of current, 4-year, and 28-year self-reported body weight in an elderly population. *Am J Epidemiol.* 1990;132:1156–1163.
38. Casey VA, Dwyer JT, Berkey CS, Coleman KA, Gardner J, Valadian I. Long-term memory of body weight and past weight satisfaction: a longitudinal follow-up study. *Am J Clin Nutr.* 1991;53:1493–1498.
39. Miller A, Springern K, Buckley L, Williams E. Diets incorporated. *Newsweek.* September 11, 1989:56–60.
40. Willett W. Challenges for public health nutrition in the 1990s. *Am J Public Health.* 1990;80:1295–1298.
41. Lissner L, Odell PM, D'Agostino RB, et al. Variability of body weight and health outcomes in the Framingham population. *N Engl J Med.* 1991;324:1839–1844.

Intestinal Parasites among North Carolina Migrant Farmworkers

Stephen D. Ciesielski, PhD, John R. Seed, PhD, Juan C. Ortiz, MSPH, and Jobe Metts, MD

ABSTRACT

Objectives. The public health significance of intestinal parasitism among migrant farmworkers has been poorly defined. We report a three-part study in which we attempted to provide population-based estimates and identify risk factors for infection.

Methods. Parasite prevalence and incidence were estimated from two cross-sectional studies and a longitudinal study. We used questionnaires and clinical measurements such as hematocrits to examine risk factors and health effects.

Results. In the 1987 convenience sample (n = 265), parasite prevalence ranged from 28% among the US born to 86% among the Central American born, with no significant difference between the prevalence of pathogens among US-born (14%) and Mexican-born (24%) subjects ($P = .12$, χ^2 test). High prevalences were also found in the 1988 random sample. An annual incidence of 9.5% (n = 74) was demonstrated in 1988. Symptoms such as abdominal pain were associated with infection, and lower hematocrits were associated with hookworm infection ($P = .02$, t test).

Conclusions. Infection with intestinal parasites appears to be an occupational hazard of farmwork, necessitating improved working and living conditions and greater clinical awareness. (*Am J Public Health.* 1992;82:1258–1262)

Introduction

Migrant farmworkers work and live in substandard hygienic conditions,[1–4] often in areas of the United States formerly hyperendemic for hookworm and other intestinal parasites. Immigrants from developing regions often enter farmwork, increasing the potential for transmission of parasites. Discussions of the health of farmworkers usually cite intestinal parasites as a major problem.[1,4,5] However, previous research has focused on children, a small minority of the population, and made little attempt to quantify transmission and risk factors or to describe associated morbidity.[2,3] We report the results of a three-part study of intestinal parasites among farmworkers in North Carolina, conducted in 1987 and 1988.

Methods

The 1987 prevalence study used a convenience sample of migrant camps in eastern North Carolina, including a representative range of camp condition and size. All individuals living in selected camps were eligible subjects. In 1988 a clustered random sample was obtained by selecting a 10% probability sample from a pooled list of camps. Because of high nonresponse rates, US-born Blacks were excluded from the 1988 sample. Our 1988 longitudinal study included all subjects relocated from the 1987 prevalence study.

After obtaining informed consent, we used pretested, verbally administered questionnaires in English, Spanish, or Haitian Creole to obtain data on demographic, health, and occupational variables. Parents supplied information for children less than 16 years old.

Subjects were given transport kits containing poly-vinyl alcohol and 10% formalin (Trend Scientific, St. Paul, Minn.). After our formalin ethyl-acetate concentration (CON-Trate System, Trend) of the samples, we employed Lugol's iodine and Tri-Chrome stain (Trend) in our microscopic examination. We used Direct Fluorescent Antibody tests (Meridian Diagnostics, Inc., Cincinnati, Ohio) to screen for *Cryptosporidium*, with acid fast dimethyl-sulfoxide stain (Trend) for confirmation. Water from camps in the 1988 prevalence and incidence studies was tested for total and fecal coliforms (membrane filtration method) using standard methods.[6]

Our statistical analysis was performed with SAS and consisted of χ^2 tests, Fisher's Exact Tests, t tests, and logistic regression. The 95% confidence intervals (CI) are not test based.

Results

1987 and 1988 Cross-sectional Studies

In 1987, 265 subjects in 24 camps in three counties participated; 54% of subjects provided one stool sample, 39% provided two, and 7% provided three. Nonresponse was 5% among Haitians, 15%

Stephen D. Ciesielski, John R. Seed, and Juan C. Ortiz are with the Department of Parasitology and Laboratory Practice, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC. Jobe Metts is with the University of North Carolina School of Medicine, Chapel Hill, NC.

Requests for reprints should be sent to Stephen D. Ciesielski, PhD, Department of Parasitology and Laboratory Practice, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

This paper was submitted to the Journal April 4, 1991, and accepted with revisions March 24, 1992.

among Central Americans, 35% among Mexicans, and 55% among US-born Blacks. The 1988 random sample included 181 subjects from 20 camps in four counties; 47% provided two samples, and 53% one. The 1988 nonresponse was similar to 1987, but higher among Central Americans (50%).

Most subjects in 1988 were young adult males from rural backgrounds with little education (Table 1). Demographics of the 1987 convenience sample were similar, except that US-born Blacks were included. Parasite prevalence was high in both years and usually varied little between years (Table 2). In 1987, 52% of all subjects were positive vs 58% in 1988. Most infections were with pathogens (defined in Table 2). In 1987 hookworm was the most prevalent pathogen among the foreign born, with a prevalence of 58% among Central Americans (28% in 1988) and 20% (24% in 1988) among Haitians (Table 2). The 1988 pathogen prevalence among both Haitians and Central Americans was significantly greater than among Mexicans ($P = .023$, relative risk (RR) = 2.7, 95% CI 1.3, 5.6; and $P = .030$, RR = 2.48, 95% CI 1.1, 5.7, χ^2 tests, respectively). One Central American subject had a *Taenia* infection.

In 1987, parasite prevalence was 28% and pathogen prevalence was 14% among the US born (50% of infected subjects were Hispanic and 50% US-born Blacks). The US-born subjects' prevalence of *Giardia* exceeded that of the Haitian born (RR = 7.2, 95% CI 1.0, 51, χ^2 test) and Mexican born (RR = 3.6, 95% CI .44, 29, χ^2 test); their *H. nana* prevalence exceeded that of the Mexican born; and their prevalence of *E. histolytica* and *E. coli* was comparable to foreign-born subjects. There was no significant difference in pathogen prevalence between the US and Mexican born ($P = .120$, χ^2 test).

In both years of the study, foreign-born subjects' return visits to their countries of origin were not associated with infection ($P > .05$, χ^2 tests). Measures of previous exposure in country of origin (age, sex, educational level, rural upbringing, previous agricultural work) were not associated with infection among foreign-born subjects in 1987 ($P > .05$, logistic regression). Among subjects in the 1988 sample, rural upbringing and previous farmwork were associated with all classes of parasite except protozoa ($P < .05$, logistic regression). Among foreign-born subjects of both years, grouped by birthplace, there were no differences in years of US residence between infected and uninfected subjects ($P = .24$–$.81$, t tests).

TABLE 1—Demographics of Subjects in 1988 Cross-sectional Study of Intestinal Parasites in Migrant Farmworkers

	All Subjects (n)[a]	Haitians (n)[a]	Hispanics (n)[a]
Male, %	80 (181)	74 (50)	83 (129)
Over 16 years of age, %	90 (181)	100 (50)	86 (129)
Mean age	29 (181)	39 (50)	25 (129)
Mean years in farmwork	4.8 (134)	5.5 (48)	3.9 (96)
Mean years in US	4.9 (145)	5.6 (48)	4.6 (97)
Mean exits from US	1.2 (145)	0.5 (48)	1.6 (97)
6 or fewer years of school, %	65 (145)	80 (48)	54 (97)
Rural upbringing, %	61 (145)	62 (48)	61 (97)
Previously in agriculture, %	38 (134)	60 (48)	24 (86)

[a]Number of subjects varies because some variables were not applicable to all subjects (particularly children).

In 1987, prevalence of intestinal parasites in Haitian camps ranged from 22 to 70%, in Mexican camps 0 to 60%, and in Black American camps 0 to 30%. This variation was not associated with years in the United States, age, gender, previous risk factors, or condition of camps.

Many subjects reported workplace conditions or behaviors with risk for infection (Table 3), but these variables were not associated with infection ($P > .05$, logistic regression). Among US-born children, day-care attendance was not associated with infection ($P = .574$, χ^2 test), but infection was associated with children accompanying parents to harvest fields ($P = .032$, RR = 2.4, 95% CI 1.3, 4.6, χ^2 test).

Reported symptoms had many significant associations with parasitic infections of different classes (Table 4). Symptoms were consistently associated with pathogenic parasites but not with nonpathogens (not shown).

The mean hematocrit of subjects (in 1988) with hookworm infection was 39.4% and that of the uninfected was 43.0%, a significant difference ($P = .02$, t test). The difference was most pronounced among Hispanic subjects (uninfected = 43.3%, infected = 37.1%, $P = .02$).

1988 Longitudinal Study

In 1988, 74 of 265 subjects (28%) were relocated in 11 camps, without nonresponse. Seven subjects (9.5%) without evidence of infection in 1987 were infected in 1988 with *G. lamblia, T. trichiura,* or *Entamoeba coli.* Latrines rather than flush toilets in camps were associated with incidence ($P = .008$, RR = 13.1, Fisher's Exact Test), as was the presence of coliforms in drinking water ($P = .011$, RR = 7.6, Fisher's Exact Test). Total coliforms were demonstrated in 44% and fecal coliforms in 26% of 27 camps (results published separately[7]).

Discussion

This study provides the first estimates of parasite prevalence among adult farmworkers, the first population-based estimates, and limited data on risk factors for infection. The results indicate that intestinal parasitism is a public health problem among farmworkers. The predominance of hookworm infection—which exceeded that of earlier studies by 6 to 20 times—is especially significant because of its pathogenic potential. However, the results also demonstrate the difficulties in applying standard epidemiologic methods to the study of migrant health. Logistical problems are partly responsible: one example is the more than 30 000 miles driven by field workers in the two years of study.

The estimates of prevalence are substantially higher than previous studies.[2,3] Although prevalence may vary among sub-populations and geographically, a population-based sampling design should provide more valid estimates. However, the majority of subjects were obtained by convenience sampling, and efforts to achieve representativeness can never avoid bias. The often close agreement in prevalence between years may suggest that the 1987 sample is reasonably representative. Nonresponse in field studies of intestinal parasites is always a problem, and the relatively high rate among some groups limits the certainty of these estimates. Nonresponse was lowest among the high-prevalence groups, which would tend to overestimate prevalence. Other

TABLE 2—Parasite Prevalence in 1987 Convenience and 1988 Random Samples of North Carolina Migrant Farmworkers (Percentage Positive)

	Birthplace									
	US		Haiti		Mexico		Central America		Total Sample	
	1987 (n = 108)	1988 (n = 11)	1987 (n = 80)	1988 (n = 50)	1987 (n = 41)	1988 (n = 87)	1987 (n = 36)	1988 (n = 32)	1987 (n = 265)	1988 (n = 181)
All parasites	28	36	66	64	56	53	86	69	52	58
All pathogens	14	0	49	52	24	29	81	50	35	37
Individual pathogens										
Hookworm	0	0	20	24	6	8	58	28	15	16
Trichuris trichiura	2	0	21	20	6	6	50	25	15	13
Ascaris lumbricoides	0	0	4	0	3	0	8	13	3	2
Strongyloides stercoralis	0	0	4	6	2	5	3	0	2	4
Taenia sp.	0	0	0	0	0	0	3	0	0.4	0
Hymenolopis nana	2	0	3	0	0	3	8	0	3	2
Giardia lamblia	8	0	1	10	3	6	11	3	6	6
Entamoeba histolytica	3	0	14	3	9	5	6	9	7	4
Cryptosporidium	0	0	2	2	2	1	3	3	2	2
Nonpathogens										
Entamoeba coli	17	36	24	12	29	26	17	25	21	22
Endolimax nana	0	0	16	10	15	8	17	9	9	8
Iodamoeba butschli	1	0	4	0	7	3	11	0	4	2
Blastocystis hominis	3	0	1	2	2	1	6	0	3	1

TABLE 3—Percentage of Subjects Reporting Risk in Occupational Factors for Transmission of Intestinal Parasites, 1987 Convenience and 1988 Random Samples

	US-born Blacks		Mexicans		Central Americans		Haitians		All Subjects	
	1987 (n = 56)	1988*	1987 (n = 58)	1988 (n = 61)	1987 (n = 31)	1988 (n = 20)	1987 (n = 69)	1988 (n = 47)	1987 (n = 214)	1988 (n = 131)
Presence of field sanitation, sometimes or never	85	...	88	88	88	75	100	55	91	75
Presence of hand-washing water in fields, sometimes or never	64	...	75	77	72	85	82	21	58	58
Availability of toilet paper in fields, sometimes or never	61	...	45	36	37	20	81	89	60	57
Common drinking cup used in fields, always or sometimes	60	...	71	75	80	95	50	53	63	69
Frequency of working without shoes (or with sandals), always or sometimes	53	...	22	20	30	55	15	8	29	30

*No Black Americans were included in the 1988 sample.

factors (such as often obtaining only one stool sample) tend towards underestimation.

Transmission of parasites in the migrant population is indicated by several results:

1. The parasite prevalence of 28% among US-born subjects in 1987. Although the source of infection cannot be determined with confidence in a cross-sectional study, it is a reasonable assumption that most infections among the US born were acquired in farmwork, especially because most children had lived in this environment their entire lives.
2. The absence of a significant difference between the prevalence of pathogenic parasites among the US and Mexican born in the 1987 sample.
3. The prevalence of several parasite species among the US-born subjects that equaled or exceeded that of the foreign born.
4. The absence of any difference in years of US residence between infected and uninfected subjects. Parasite prevalence among nonfarmworker immigrants to the United

TABLE 4—Reported Symptoms and Infection with Intestinal Parasites

Year of Study	Parasite	Group	Symptom	P*	RR	95% CI
1987	Pathogens	All subjects	Nausea	.05	2.2	1.0, 4.8
		All subjects	Appetite loss	.002	3.2	1.7, 5.9
1987	Hookworm	All subjects	Appetite loss	.021	2.5	1.3, 4.9
		All subjects	Abdominal pain	.034	2.6	1.5, 5.0
1988	All parasites	All subjects	Abdominal pain	.0009	3.5	2.9, 4.0
		Hispanics	Abdominal pain	.017	2.7	2.0, 3.9
1988	Pathogens	All subjects	Abdominal pain	.0009	2.8	2.2, 3.4
		Haitians	Abdominal pain	.014	2.6	1.5, 3.8
		Hispanics	Abdominal pain	.002	3.0	1.9, 4.3

*From chi-square tests.

States decreases with increasing years of US residence.[8]

5. The annual incidence rate of 9.5%.

Use of a control group could have demonstrated more conclusively that the prevalence among US-born farmworkers exceeds the nonmigrant population. However, only longitudinal study can prove that infections are acquired in farmwork. Identification of a suitable control group would be difficult; in North Carolina there is no demographically similar group of working adults with such a low socioeconomic status. Future studies should employ one if conducted in Florida or California, where demographics make construction of a control group more feasible.

A cross-sectional design limits assessment of risk factors. Permission to directly inspect workplace conditions was denied; subjects reported conditions for the month prior to study, but working conditions may change weekly or daily. A sample size several times larger than described here would be needed to demonstrate an association between infection and risk factors cross-sectionally. Nevertheless, the potential for transmission is evident in the lack of field sanitation and other workplace conditions. The association between infection and children accompanying parents to the fields indicates occupational risk.

Obstacles to conducting longitudinal study are also substantial, given the transience of farmworkers. The incidence rate of 9.5% reported here is the first attempt to quantify transmission. However, problems include the small sample size (74 subjects) and the possibility of false negatives in 1987. The associations between infections counted as incidence and camp conditions such as latrines and coliforms in drinking water support the classification of these as new infections. However, because all three incident parasites (especially *Trichuris*) can be transmitted via nonwater routes, the importance of water as a means of transmission remains undefined.

The clinical significance of intestinal parasitism among farmworkers has been inconclusively addressed in previous studies. In this study, hematocrits and reported symptoms indicate that parasitism causes morbidity among farmworkers, although many infections may be asymptomatic. Although only one *Taenia* infection was demonstrated, it suggests the possible spread of cysticercosis among farmworkers. The seroprevalence of cysticercosis was 10% in a random sample of NC farmworkers.[9]

Descriptions of the health of farmworkers usually cite intestinal parasites as a paramount problem. Yet no one would maintain that parasitic infections are of greater significance than hypertension, diabetes, or the human immunodeficiency virus. The shock value of parasites is obviously one reason for this emphasis. But it is also true that parasitic infections are a sentinel health event. Their transmission indicates that working and living conditions place workers at risk for a variety of other enteric infectious diseases—both bacterial and viral—transmitted in a similar fashion. Farmworkers have rates of diarrhea 20 times that of the urban poor,[4] and such pathogens must contribute to this disparity.

We conclude that there is good evidence for the transmission of intestinal parasites within the conditions of migrant farmwork, and that infection with parasites is associated with some degree of morbidity. Workplace conditions present clear potential sources of infection. Farmworkers and their families are obviously at greatest risk. However, parasitic ova and cysts remain viable for long periods of time on produce,[10–12] creating some risk for the general public. In 1991 a North Carolina man who lived on a family farm employing migrant farmworkers (from areas endemic for cysticercosis) was diagnosed as having cerebral cysticercosis, a potentially fatal disease.[13] This infection was almost certainly acquired from the farm's employees. On the basis of the results of this study, we make the following recommendations:

- Enforce more strictly the regulations concerning field sanitation and hygiene.
- Replace latrines in migrant camps with flush toilets.
- Improve the surveillance of the microbiologic quality of drinking water in migrant camps.
- Routinely include intestinal parasitism on differential diagnoses of gastrointestinal complaints of all groups of farmworkers, including US-born blacks.
- Routinely screen high-risk patients such as Haitians and Central Americans for intestinal parasites, even when they are presenting other complaints: treatment of asymptomatic individuals can have an important public health function. □

Acknowledgments

Funding for this study was provided by the Milbank Memorial Fund, New York, NY, the Rural Community Assistance Program, Leesburg, Va, and the Tri-County Community Health Center, Newton Grove, NC.

We thank Nora Selva-Sutter for her assistance in diagnostic microscopy and Gustavo Saldias, MPH, Kathy Sarri, MPH, Roland Dorancy, Michael Baker (TCCHC director), Cynthia Behler, RN, and Kathleen Stanley (director RCAP). Gratitude is also expressed to

Edward Wrenn, MD, Douglas Esposito, MD, MPH, Edwin Dimi Stephen, and Mary Sween.

References

1. Arbab DM, Weidner B. Infectious disease and field water supply and sanitation among migrant farm workers. *Am J Public Health.* 1986;76:694–695.
2. Ortiz J. The prevalence of intestinal parasites in Puerto Rican farm workers in Western Massachusetts. *Am J Public Health.* 1980;74:1103–1105.
3. Ungar B, Iscoe E, Cutler J, Bartlett JG. Intestinal parasites in a migrant farmworker population. *Arch Intern Med.* 1986; 146:513–515.
4. Wilk VA. *The Occupational Health of Migrant and Seasonal Farmworkers in the United States.* Washington, DC: Farmworker Justice Fund Inc.; 1986.
5. Pendergrass JA. Field sanitation; final rule. *Federal Register.* May 1, 1987;52:16050–16096.
6. Clesceri LS, Greenberg AE, Trussell RR, eds. *Standard Methods for the Examination of Water and Wastewater.* Washington, DC: American Public Health Association; 1981.
7. Ciesielski S, Handzel T, Sobsey M. The microbiologic quality of drinking water in North Carolina migrant labor camps. *Am J Public Health.* 1991;81:762–764.
8. Winsberg GR, Sonnenschein E, Dyer AR, Schnadig V, Bonilla E. Prevalence of intestinal parasites in Latino residents of Chicago. *Am J Epidemiol.* 1975;102:526–532.
9. Ciesielski SD, Estrada J, Seed JR, Wrenn E. The seroepidemiology of tropical disease among North Carolina migrant farmworkers. Unpublished manuscript.
10. Geldreich EE, Bordner RH. Fecal contamination of fruits and vegetables during cultivation and processing for market: a review. *J Milk Food Tech.* 1971;34:184–195.
11. Rudolfs W, Falk LL, Ragotzkie RA. Contamination of vegetables grown in polluted soil. II. Field and Laboratory studies on *Entamoeba* cysts. *Sewage Ind Waste.* 1951;23:478–485.
12. Rudolfs W, Falk LL, Ragotzkie RA. Contamination of vegetables grown in polluted soil. III. Field studies on *Ascaris* eggs. *Sewage Ind Waste.* 1951;23:656–660.
13. Centers for Disease Control. Locally acquired neurocysticercosis—North Carolina, Massachusetts, and South Carolina, 1989-1991. *MMWR.* 1992;41:1–4.

Seventh National Conference on Chronic Disease Prevention and Control

The Centers for Disease Control, the Association of State and Territorial Health Officials, and the Association of State and Territorial Chronic Disease Program Directors will cosponsor the Seventh National Conference on Chronic Disease Prevention and Control, October 21 to 23, 1992, in Salt Lake City, Utah. The conference is open to the public. This year's theme is "Prevention: Can It Be Marketed, Managed, and Make a Difference?"

The conference will emphasize interactions among federal, state, and local health departments, voluntary health agencies, and professional organizations. Presentations will be given in plenary, invited, abstract, poster, and workshop sessions. Topics will include: "Universal Access: Future Implications for Prevention and Public Health"; "Behavioral Research: What Have We Applied?"; "Prevention Effectiveness: Documenting Our Successes"; "The Tobacco Industry's Tactics and Their Implications for Public Health"; "The Family High Risk Program: The Ultimate in Targeted Interventions"; "Clinic-Based Interventions"; "Reaching Children and Adolescents"; "Smoking and Health"; "Nutrition and Chronic Disease Prevention"; "Reaching the Underserved"; "Using the Latest in Computer Software for Public Health"; "Enlisting Minority Health Organizations in Chronic Disease Advocacy"; and "International Chronic Disease Prevention and Control."

Additional information is available from Mr. Jack Friel, National Center for Chronic Disease Prevention and Health Promotion, Mailstop K-43, Centers for Disease Control, Atlanta, GA 30333; tel. (404) 488-5390, fax (404) 488-5962.

Restricted Activity Days among Older Adults

ABSTRACT

Objectives. The number of restricted activity days experienced by an individual in the course of a year is an important measure of functional well-being, particularly for older adults. We sought to determine multivariate associations between restricted activity days and various health conditions.

Methods. We used data from the 1984 Supplement on Aging of the National Health Interview Survey to estimate the relationship between restricted activity days and age, gender, and the presence or absence of selected chronic conditions and falls for all noninstitutionalized people aged 65 years and over. Chronic conditions and falls accounted for most of the variance in the model.

Results. Of an annual average of 31 restricted activity days, 6 days were associated with falls; 4 days with heart disease; 4 days with arthritis and rheumatism; 2 days each with high blood pressure, cerebrovascular disease, and visual impairment; and 1 day each with atherosclerosis, diabetes, major malignancies, and osteoporosis.

Conclusions. These results can be used in estimating the potential impact of health promotion programs on the health status of noninstitutionalized older adults. (*Am J Public Health.* 1992;82:1263–1267)

Michael R. Kosorok, PhD, Gilbert S. Omenn, MD, PhD, Paula Diehr, PhD, Thomas D. Koepsell, MD, MPH, and Donald L. Patrick, PhD, MSPH

Introduction

A restricted activity day, as defined in the National Health Interview Survey, is a day in which an individual spends over half of the day in bed, home from work or school, or cutting down on usual activities because of illness or injury.[1] The number of restricted activity days experienced by an individual in a year is an important measure of functional status and well-being,[2] assessed in the National Health Interview Survey for many years.[3] One of the surgeon general's two major goals for older adults, as stated a decade ago in *Healthy People,*[4] was to improve the health and quality of life by reducing the average annual number of days of restricted activity due to acute and chronic conditions by 20%, to fewer than 30 days per year for people 65 years of age and older. (A change in the order and wording of the 1982 National Health Interview Survey questions regarding restricted activity days resulted in about 20% fewer reported restricted activity days per person.[3,5] Thus, if we adjust for the questionnaire change, the goal of 30 days [as written before 1982] is probably equivalent to 24 days, as measured by post-1982 questionnaires [30 days less 20%].)

Because the meaning of "restricted activity" depends in part on what activities are usual for an individual and on his or her subjective judgment about degree of limitation due to injury or illness, the restricted activity day is thought by some to be "less than an ideal indicator of the health status of older adults."[6] However, Scholes et al. recently demonstrated that restricted activity days correlate well with important physical health status measures.[7]

The purpose of this paper is to estimate the relationship between restricted activity and the presence or absence of major chronic conditions, falls, increasing age, and gender. Several published studies have shown bivariate associations between the number of restricted activity days per person per year and various health conditions.[1,8,9] In addition, Verbrugge et al. investigated multivariate associations (comorbidity) between several chronic conditions and physical well-being as measured by the number of activities of daily living performed with difficulty.[10] We present multivariate analyses of the association of health conditions with restricted activity days. Such information may be useful in identifying risk factors for disability, assessing priorities for prevention among older adults, and predicting the potential population impact of health promotion interventions.

Methods

Data and Selection of Variables

Our data set was drawn from the 1984 Supplement On Aging of the National Health Interview Survey, based on a ran-

Michael R. Kosorok is with the Departments of Biostatistics and Statistics at the University of Wisconsin, Madison. Gilbert S. Omenn is with the School of Public Health and Community Medicine and the Center for Health Promotion in Older Adults, Paula Diehr is with the Departments of Biostatistics and Health Services, Thomas D. Koepsell is with the Departments of Epidemiology and Health Services, and Donald L. Patrick is with the Department of Health Services, all at the University of Washington, Seattle.

Requests for reprints should be sent to Gilbert S. Omenn, MD, PhD, School of Public Health and Community Medicine SC-30, University of Washington, Seattle, WA 98195.

This paper was submitted to the Journal October 22, 1991, and accepted with revisions March 3, 1992.

TABLE 1—Health Conditions Selected for Analysis

Condition	Survey Question Time Frame	Condition Prevalence per 1000	Restricted Activity Days: Average per Person with Condition	Restricted Activity Days: Percentage of Total
Visual impairment	Now	112.8	60.2	21.6
Cataracts	Now	176.6	40.6	22.9
Glaucoma	Now	42.4	45.1	6.1
Hearing impairment	Now	261.3	36.9	30.7
Tinnitis	Now	123.0	39.9	15.6
Falls	Past 12 months	203.7	62.1	40.4
Osteoporosis	Ever	33.1	71.8	7.6
Hip fracture	Ever	35.8	55.2	6.3
Atherosclerosis	Ever	105.5	64.1	21.5
High blood pressure	Ever	440.5	39.9	56.1
Ischemic heart disease	Ever	144.9	57.9	26.7
Other heart disease	Ever	89.1	68.6	19.5
Cerebrovascular disease	Ever	66.1	83.6	17.6
Diabetes	Past 12 months	97.8	54.4	17.0
Varicose veins, lower extremities	Past 12 months	95.4	42.3	12.9
Major malignancy	Ever	52.2	56.4	9.4
Arthritis and rheumatism	Past 12 months	495.2	39.4	62.2
Overall average	...	...	31.4	...

dom sample of adults 55 years of age or older from the noninstitutionalized US population during 1984.[11,12] We have examined the findings for the 11 497 subjects 65 years of age or older. Information on a variety of health-related topics was obtained for each individual, including the number of restricted activity days experienced in the 2-week period preceding the interview and the presence or absence of selected health conditions from a uniform list of 43 conditions.[12] The Supplement on Aging questionnaire differs from the standard National Health Interview Survey practice that randomly assigns only one of six different lists of conditions to each subject. The Supplement on Aging data also were used by Verbrugge et al. in their research on comorbidity.[10] The 14-day data were converted to annual rates.

In the interest of parsimony, clinically similar chronic conditions were combined prior to development of the multivariate models; conditions associated with less than 5% of the total annual restricted activity days were excluded. Age was recorded as a binary variable (0 = age 65 through 74, 1 = age 75 or over). Table 1 shows the average annual restricted activity days overall (31.4 days per year) for people who reported each health condition analyzed. For example, people with visual impairment averaged 60.2 restricted activity days per year and accounted for 21.6% of all annual restricted activity days.

These prevalence rates are based on self-reports by noninstitutionalized respondents. Alzheimer's disease is an obvious example of underreporting. Other weaknesses inherent in the Supplement on Aging data set that limit its ability to provide information on components of morbidity among institutionalized adults have been discussed extensively by Verbrugge et al. However, these authors concluded that if the scope of the research is to determine how disabling certain reported chronic conditions are for noninstitutionalized community residents, then the data set is appropriate and advantageous because of its size and national span.[10]

The National Center for Health Statistics has compared interview data with medical records data from a population in the Kaiser Foundation Health Plan of Southern California and reported good agreement, especially for cardiovascular disease.[13] Unfortunately, several additional important health conditions were not adequately addressed in the Supplement on Aging questionnaire and could not be included in the model, including chronic lung disease, back problems, chronic gastrointestinal disorders, paralysis, other problems with lower extremities, other musculoskeletal disorders, and depression.

Preliminary Analysis

Because the National Health Interview Survey is a stratified, clustered random sample and not a simple random sample, the analyses must take into account the survey design effect.[12] We performed preliminary analyses using ordinary weighted least squares and then estimated the final models (including variances and *P* values) using the stratified, clustered sample procedures found in PC-CARP software.[14]

We also analyzed bed disability days but preferred restricted activity days both because of the use of this variable by the surgeon general[4] and because our analyses of both variables showed restricted activity days (after correction for the 1982 National Center for Health Statistics/National Health Interview Survey questionnaire change) to be a more stable dependent variable than bed disability days, as in the analyses of Scholes et al.[7]

Final Analysis

With the number of restricted activity days experienced in a 2-week period as the dependent variable, the independent variables selected for analysis included age, sex, and presence or absence of the 17 selected health conditions during the time periods specified in Table 1. We performed bivariate regression of restricted activity days on each independent variable in order to estimate zero-order associations with restricted activity days. We performed multiple stepwise linear regression for a model in which all independent main effects were forced to enter, and then used an entry *P* value of .05 and an exit *P* value of .10 in order to arrive at a parsimonious model containing main effects. Leaving these main effects in, we used the same stepwise regression methods to search for significant interactions between each health condition and either age or sex and between age and sex. Leaving these main effects and significant interactions in, we used the same stepwise procedures to examine any significant two-way interactions among the set of 59 health condition interactions that were associated with at least 5% of the total annual restricted activity days experienced (same lower bound as in the initial selection of independent variables). Each of the above models was confirmed in PC-CARP.

Results

The relationships between main effects (independent variables) and restricted activity days are summarized in Table 2. Since the 2-week averages result from properly weighted observations taken from all quarters of the year, this conversion to annual rates should produce properly weighted averages for all seasons.[11] The second column in Table 2 contains the regression coefficients for each variable when that was the only variable in the model. This value estimates the differ-

ence between populations with the specified condition and populations without. For age, the condition is being 75 years old or older; for sex, female. In this column, cerebrovascular disease shows the greatest association with restricted activity (55.9 days).

The third column in Table 2 shows coefficients for the final main effects model. A zero without any decimals indicates that the variable was left out of the model by the model selection criteria and its effect was thus taken to be zero. In this column, the effect of cerebrovascular disease is still the highest (35.2 days), but the magnitude of the coefficient is lower, because some of the apparent effect of cerebrovascular disease can be accounted for by the presence of other conditions. In fact, all coefficients in the final main effects model are smaller than their zero-order counterparts, indicating that many of these conditions are correlated.

The fourth column of Table 2 presents estimates with an interpretation similar to those of the third column, except that all variables were forced to enter into the model. By comparing this column with the final main effects model (third column), the stability of the coefficients of variables left in the final main effects model can be verified. The variables left out of the final main effects model do not have a large influence on the variables that were left in.

When the age by sex interaction and all interactions between each health condition and either age or sex were examined as possible additions to the final main effects model, none was statistically significant. When all 59 two-way interactions among selected health conditions were examined in the presence of the final main effects model, only four remained significant at the .05 level in PC-CARP. The resulting model will hereafter be referred to as the final interaction model. The estimated main effect and interaction coefficients for this model are given in the fifth column of Table 2, along with the model R^2 and overall P value. Confidence intervals for the parameters of this model are not given because it is difficult to interpret confidence intervals in the presence of interactions.

Main effect coefficients that differ between the final main effects model and the final interaction model are part of interactions (e.g., "falls by arthritis and rheumatism"). In addition, the baseline estimate increases in the interaction model. Overlap between main effects cannot be accounted for in a main effects model, resulting in a low estimate of the effect of conditions not included in the model. The actual average from the data for people with none of the 17 selected conditions is 10.5 restricted activity days per year. The estimated baseline level of 8.4 in the final interaction model is closer than the estimate from the final main effects model (6.2).

TABLE 2—Estimated Model Coefficients for Four Different Models Predicting the Annual Number of Restricted Activity Days

Final Condition (Independent Variable)	Regression Coefficients (In Days per Annum) Predicted by Four Models: Zero Order	All Main Effects (95% Confidence Interval)	Effects Forced in	Final Interaction
Visual impairment	32.5	16.3 (9.4, 23.2)	15.7	16.6
Cataracts	11.2	0 ...	0.0[a]	0
Glaucoma	14.3	0 ...	5.9[a]	0
Hearing impairment	7.4	0 ...	−1.4[a]	0
Tinnitis	9.7	0 ...	1.8[a]	0
Falls	38.6	27.4 (23.1, 31.8)	27.3	19.7
Osteoporosis	41.8	27.1 (16.2, 37.9)	27.2	26.8
Hip fracture	24.7	13.0 (1.4, 24.6)	12.5	13.3
Atherosclerosis	36.6	13.6 (5.6, 21.7)	13.2	7.9
High blood pressure	15.3	4.5 (1.4, 7.6)	4.5	3.2
Ischemic heart disease	31.0	15.0 (9.2, 20.8)	15.0	15.1
Other heart disease	40.9	25.2 (16.0, 34.5)	25.1	22.0
Cerebrovascular disease	55.9	35.2 (25.8, 44.7)	5.0	34.9
Diabetes	25.5	14.8 (7.8, 21.7)	14.7	8.5
Varicose veins, lower extremities	12.1	0 ...	2.7[a]	0
Major malignancy	26.4	21.9 (12.8, 30.9)	21.8	21.8
Arthritis and rheumatism	15.9	7.4 (4.0, 10.8)	7.3	4.8
Age (75 years or older)	9.0	0 ...	2.1[a]	0
Sex (female)	3.9[a]	0 ...	−0.9[a]	0
Falls × arthritis and rheumatism				13.6
Atherosclerosis × diabetes				38.6
High blood pressure × other heart disease				18.3
Other heart disease × hearing impairment				−20.4
Overall average	31.4	31.4	31.4	31.4
Baseline level	...	6.2 (3.8, 8.5)	5.8	8.4
Regression R^2	...	.08 ...	.08	.08
Overall P value	...	.001 ...	.001	.001

Note. The coefficient values for each model can be interpreted as the additional number of days experienced in association with the presence of the corresponding condition (n = 11 497 persons 65 years of age and older).
[a]Not significant at the .05 level.

Using the final main effects model, we estimated the average number of restricted activity days per year for persons 65 years old or older that may be directly associated with each condition. These estimates (Table 3) were calculated by multiplying the model coefficient for the condition by the prevalence of the condition per individual, that is, column 3 = column 3 in Table 2 × (column 2 in Table 3/1000). Column 4 in Table 3, in which the resulting measure is expressed as a percentage of the overall average number of restricted activity days, is analogous to the "etiologic fraction" sometimes used as a measure of effect in epidemiologic studies.[15] It is interesting to note that whereas cerebrovascular disease has the largest model coefficient, both falls and arthritis, when multiplied by prevalence, become larger contributors to total annual restricted activity days.

Values for the final interaction model were calculated in the same manner. For those main effects that were part of an interaction, the corresponding days and percentage of days for the associated interaction were added to the main effect values (columns 5 and 6, Table 3). For example, the average annual days directly associated with falls was calculated by multiplying its prevalence by its coefficient [(203.7/1000) × 19.7 = 4.0] and adding this product to the prevalence of the arthritis and rheumatism combination multiplied by its coefficient {4.0 + [(123.7/1000) × 13.6] = 5.7}.

TABLE 3—Estimates of Average Annual Restricted Activity Days per Person Directly Associated with Conditions Significant in the Final Main Effects Model and in the Final Interaction Model

Condition	Condition Prevalence per 1000	Final Main Effects Model: Annual Days Directly Associated with Condition	Final Main Effects Model: Days Directly Associated with Condition, %	Final Interaction Model: Annual Days Directly Associated with Condition	Final Interaction Model: Days Directly Associated with Condition, %
Visual impairment	112.8	1.8	5.9	1.9	6.0
Falls	203.7	5.6	17.8	5.7	18.1
Osteoporosis	33.1	0.9	2.9	0.9	2.8
Hip fracture	35.8	0.5	1.5	0.5	1.5
Atherosclerosis	105.5	1.4	4.6	1.5	4.7
High blood pressure	440.5	2.0	6.3	2.3	7.4
Ischemic heart disease	144.9	2.2	6.9	2.2	7.0
Other heart disease	89.1	2.2	7.2	2.3	7.2
Cerebrovascular Disease	66.1	2.3	7.4	2.3	7.4
Diabetes	97.8	1.4	4.6	1.5	4.7
Major malignancy	52.2	1.1	3.6	1.1	3.6
Arthritis, rheumatism	495.2	3.7	11.7	4.1	13.0
Falls × arthritis and rheumatism	123.7	...	...	1.7	5.4
Atherosclerosis × diabetes	16.4	...	...	0.6	2.0
High blood pressure × other heart disease	50.4	...	...	0.9	2.9
Other heart disease × hearing impairment	30.5	...	...	0	0
Overall average	...	31.4	100.0	31.4	100.0
Baseline level	1000.0	6.2	19.6	8.4	26.9

Some of the days directly associated with a condition are counted twice. For example, the days directly associated with the falls by arthritis and rheumatism interaction are added to both the falls main effect and the arthritis and rheumatism main effect. The overlap resulting from each interaction (the value counted twice) is given in the last four lines of columns 5 and 6 in Table 3. Summing the percentages in the last column for the baseline level and main effects and then subtracting the overlap values from the interactions will result in 100% (within rounding error).

Discussion

Neither age, sex, nor any interaction with age or sex was significantly associated with restricted activity days in the presence of self-reported conditions. Verbrugge et al. obtained similar results from their regression analyses.[10] This finding supports the concept that the functional decline commonly associated with age is not a direct consequence of advancing age per se.[16,17] Instead, functional impairment may be viewed as resulting from certain medical conditions associated with age, and against which preventive interventions may be aimed.

These results also may be helpful in prioritizing health conditions for interventions: the larger the number of restricted activity days attributable to a given health condition, the greater the potential overall impact of preventing that condition.

Subject to certain assumptions discussed below, the estimates from the final interaction model (Table 3) can be used to estimate the potential population impact of specific health interventions for noninstitutionalized people 65 years of age or older. For example, suppose we have an intervention against falls that may reduce by 50% the total number of people experiencing a fall in 1 year. From Table 3, we know that falls are associated with about 18% of all restricted activity days experienced in the United States; we would, therefore, estimate that the proposed intervention could reduce the total number of restricted activity days by about 9%.

The effects of interventions with multiple benefits, such as those against smoking and/or high blood pressure, could be estimated similarly by determining the interrelationships between the various risk factors and health conditions involved and then multiplying the resulting reductions in health condition prevalences by the direct effect estimates found in Table 3. Care must be taken to ensure that the estimated overlap is subtracted. For example, a 50% reduction in both falls and arthritis and rheumatism would result in an estimate of (50% × 18.1%) + (50% × 13.0%) − (50% × 5.4%) = 12.9% reduction, where 5.4% is the overlap in percentage of days associated with the falls by arthritis and rheumatism interaction. The resulting value could be used as an estimate of the potential effectiveness of the given intervention in reducing restricted activity for noninstitutionalized populations 65 years of age and older.

The R^2 values at the bottom of Table 2 indicate that only about 8% of the variability in restricted activity days was explained by the conditions in the models. This is not surprising, since people have restricted activity days for many other reasons than falls and chronic conditions. Thus, this model would not be appropriate for predicting the number of restricted activity days for an individual. However, it provides some useful guides for patient groups, as can be seen in the confidence intervals in Table 2 for each parameter in the final main effects model. The effects of chronic conditions on restricted activity days are statistically significant, large, and consistent with clinical impressions. For example, the impact of arthritis on restricted activity days for people who suffer from arthritis is smaller than the impact of cerebrovascular disease for people who have that disease, although both conditions are significantly associated with restricted activity days and arthritis is much more prevalent. There is a larger potential for reduction in restricted activity days for the cerebrovascular disease group.

We cannot be certain that any given condition (falls, for example) *caused* the restricted activity days with which it was associated in this data set, or, more important, that preventing half of all falls—if we could—would reduce the average number of restricted activity days for the entire older adult population by 2.8 days per year (Table 3). The occurrence of a fall may reflect underlying frailty that itself increases the number of restricted activity days by other mechanisms. In addition, if other significant variables, such as severity or duration, were included in the model, the coefficient for falls from the final main effects model might drop below 27.4 days per person per year.

Our analysis was limited to conditions prevalent in noninstitutionalized persons and about which questions were asked in the National Health Interview Survey Supplement on Aging. Omission

of other important conditions, such as respiratory disease, depression, alcoholism, or Alzheimer's disease, could have either or both of two effects. First, if an omitted condition is associated with another condition included on our list, then we may incorrectly ascribe an effect to an included condition, when in fact part of that effect is due to the omitted condition(s) with which it is associated. Second, absence of data on important omitted conditions may show up as a low R^2 value: part of the unexplained variability in restricted activity days may have been explainable with additional predictor variables. We also do not know how direct the association is between a chronic condition and the number of restricted activity days; perhaps some chronic conditions are associated with an increase in acute conditions that, in turn, accounts for additional days of restricted activity. Nevertheless, our multivariate model represents a considerable improvement over bivariate analyses, as illustrated by the difference between the estimates of 27.4 days associated with falls in Table 2 and 62.1 days in Table 1.

Although additional data and more research are needed to increase its precision and broaden its applicability, the model presented in this paper should be useful in designing and evaluating health promotion interventions for older adults. □

References

1. Riess PW. *Current estimates from the National Health Interview Survey: United States, 1984.* Hyattsville, Md: National Center for Health Statistics; 1986. DHHS publication PHS 86-1584. (Vital and health statistics; series 10; no. 156).
2. Patrick DL, Stein J, Porta M, Porter CQ, Ricketts TC. Poverty, health services, and health status in rural America. *Milbank Q.* 1988;66:105–136.
3. Golden PM. *Health, United States, 1986* (prevention profile). Washington, DC: US Government Printing Office, 1986. DHHS publication PHS 87-1232.
4. *Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention.* Washington DC: US Dept of Health, Education, and Welfare; 1979. DHEW publication PHS 79-55071.
5. Kovar MG, Poe GS. *The National Health Interview Survey Design, 1973–84, and Procedure, 1975–83.* Hyattsville, Md: National Center for Health Statistics; 1985. DHHS publication PHS 85-1320. (Vital and health statistics; series 1; no. 18).
6. National Center for Health Statistics. *Prevention Profile. Health, United States, 1989.* Hyattsville, Md: Public Health Service; 1990.
7. Scholes D, LaCroix AZ, Wagner EH, Grothaus LC, Hecht JA. Tracking progress towards national health objectives in the elderly: what do restricted activity days signify? *Am J Public Health.* 1991;81:485–488.
8. Moss AJ, Parsons VL. *Current Estimates from the National Health Interview Survey: United States, 1985.* Hyattsville, Md: National Center for Health Statistics; 1986. DHHS publication PHS 86-1588. (Vital and health statistics; series 10; no. 160).
9. Verbrugge LM. Recent, present, and future health of American adults. *Annu Rev Public Health.* 1989;10:333–361.
10. Verbrugge LM, Lepkowski JM, Imanaka Y. Comorbidity and its impact on disability. *Milbank Q.* 1989;67:450–484.
11. National Center for Health Statistics. *Supplement on Aging to the 1984 National Health Interview Survey, Public Use Data Tapes.* Hyattsville, Md: Division of Health Interview Statistics, National Center for Health Statistics.
12. Fitti JE, Kovar MG. *The Supplement on Aging to the 1984 National Health Interview Survey.* Hyattsville, Md: National Center for Health Statistics; 1987. DHHS publication PHS 87-1323. (Vital and health statistics; series 1; no. 21).
13. Madow WG. *Net Differences in Interview Data on Chronic Conditions and Information Derived from Medical Records.* Rockville, Md: US Dept of Health, Education, and Welfare; 1973. DHEW publication HSM 73-1331.
14. Fuller WA, Kennedy W, Schnell D, Sullivan G, Park HJ. *PC CARP.* Ames, Iowa: Statistical Laboratory, Iowa State University; 1986.
15. Miettinen OS. *Theoretical Epidemiology.* New York, NY: Wiley; 1985.
16. Rowe JW, Kahn R. Human aging: usual and successful. *Science.* 1987;237:143–149.
17. Omenn GS. Prevention and the elderly: what are appropriate policies? *Health Aff.* 1990;9:80–93.

Public Health Briefs

Predictors of Smoking Cessation and Relapse in Older Adults

Marcel E. Salive, MD, MPH, Joan Cornoni-Huntley, PhD, Andrea Z. LaCroix, PhD, Adrian M. Ostfeld, MD, Robert B. Wallace, MD, and Charles H. Hennekens, MD, DrPH

ABSTRACT

We examined longitudinal changes in smoking behavior among older adults in three community cohorts of the Established Populations for Epidemiologic Studies of the Elderly. Smoking prevalence declined from 15% at baseline to 9% during 6 years of follow-up. Annual smoking cessation and relapse rates were 10% and less than 1%, respectively. Interval diagnosis of myocardial infarction, stroke, or cancer increased subsequent smoking cessation but not relapse. Although smoking cessation around diagnosis is increased, primary prevention could yield greater benefits. (*Am J Public Health.* 1992; 82:1268–1271)

Introduction

Cigarette smoking is the leading cause of premature mortality among older persons, primarily due to cancer and cardiovascular disease.[1–4] Smoking cessation has health benefits, largely cardiovascular, before and after age 65 years.[3–6] Although 4.5 million older adults continue to smoke regularly, 10 million have quit.[7] Older smokers are more likely to quit than younger smokers,[8–12] possibly as a result of disease events, severity of illness, or hospitalization.[6,7,12–13]

Previous reports from the Established Populations for the Epidemiologic Studies of the Elderly (EPESE) compared current, former, and never smokers at baseline[14] and their 5-year mortality during follow-up.[4] This study examined smoking cessation and relapse during follow-up and their predictors, particularly chronic disease events, in three community cohorts. We examined heart attack, stroke, and cancer because they are associated with smoking and increase with age.

Methods

The data were taken from the East Boston, New Haven, and Iowa and Washington counties EPESE communities, which enrolled approximately 10 300 participants 65 years of age or older beginning in 1981. Details of the methods have been published.[4,14–16] Trained interviewers conducted baseline household surveys and six annual follow-up interviews (telephone and in person) that included information on demographics, medical history, and health-related behaviors. Of eligible community residents, 80% to 84% completed initial interviews, and annual follow-up rates were above 95% in each community.

This analysis followed baseline current and former cigarette smokers to determine smoking status 3 and 6 years later. Smoking status, determined by two questions ("Do you smoke cigarettes [regularly] now?" and "On the average, how many cigarettes per day do you [usually] smoke?"), was unavailable in less than 1% of completed in-home interviews. Cessation and relapse were defined as a change in smoking status at follow-up among current and former smokers, respectively. Additional baseline questions included prior smoking ("Did you ever smoke cigarettes [regularly]?"), and age of initiation or cessation, if applicable. Number of years of smoking was computed.

All interviews included questions about the diagnosis of myocardial infarction, stroke, and cancer by a physician (or,

At the time of the study, Marcel E. Salive and Joan Cornoni-Huntley were with the Epidemiology, Demography and Biometry Program, National Institute on Aging, Bethesda, Md. Andrea Z. LaCroix is with the Center for Health Studies, Group Health Cooperative of Puget Sound, Seattle, Wash. Adrian M. Ostfeld is with the Department of Epidemiology, Yale University School of Medicine, New Haven, Conn. Robert B. Wallace is with the Department of Preventive Medicine and Environmental Health, University of Iowa, Iowa City. Charles H. Hennekens is with the Departments of Medicine and Preventive Medicine, Harvard Medical School and Brigham and Women's Hospital, Boston, Mass.

Requests for reprints should be sent to Marcel E. Salive, MD, MPH, Epidemiology, Demography and Biometry Program, National Institute on Aging, National Institutes of Health, 7201 Wisconsin Ave, Gateway Bldg, Suite 3C-309, Bethesda, MD 20892.

This paper was submitted to the Journal October 9, 1991, and accepted with revisions May 5, 1992.

Editor's Note. See related editorial by Giovino et al. (p 1203) in this issue.

in East Boston, another health care provider). Exertional chest and leg pain were assessed at baseline.[17]

Cessation and relapse rates were age adjusted by the direct method using the combined baseline population. The communities were grouped to examine factors associated with smoking cessation and relapse using categorical methods. Logistic regression models[18] used either smoking cessation or relapse as the dependent variable in multivariate analysis to adjust for demographics and smoking intensity.

Results

Smoking prevalence declined from 15% at baseline to 12% and 9% at years 3 and 6, respectively. After 3 years, 357 (28%) of 1259 baseline smokers had quit. By 6 years, the cessation rate increased to 46%; 433 of 944 baseline smokers had quit. Age-adjusted cessation and relapse rates differed little by gender and site (Table 1). Eighty-seven percent of surviving 3-year quitters remained nonsmokers at the final follow-up.

Smoking cessation in 3 years (Table 2) was independently and significantly ($P < .05$) associated with older age; never having been married; smoking fewer cigarettes per day; fewer years of smoking; no history of myocardial infarction, stroke, cancer, or report of angina or intermittent claudication at baseline; and interval diagnosis of myocardial infarction, stroke, or cancer. Smokers in Iowa were more likely to quit than those in New Haven (relative odds = 1.5, 95% CI = 1.0, 2.3), adjusting for variables in Table 2. Smoking cessation in 6 years was independently and significantly associated with non-Black race, smoking fewer cigarettes per day, and interval diagnosis of myocardial infarction, stroke, or cancer (Table 3).

Among former smokers, the relapse rate was 3%: 68 of 2050 baseline former smokers relapsed by 3 years and 44 of 1515 relapsed by 6 years, respectively. Three-year smoking relapse was predicted by younger age, heavier smoking history, and quitting smoking later in life (Table 4). Two fifths of former smokers who relapsed by year 3 had quit again by year 6, thus yielding too few persons for multivariate analysis.

TABLE 1—Three- and Six-Year Age-Adjusted Rates (Percentage during the Interval) of Smoking Cessation and Relapse, by Site and Gender

	Site		
	East Boston	New Haven	Iowa
Cessation rate[a]			
Men			
3-year	34	28	33
6-year	47	47	61
Women			
3-year	35	26	43
6-year	41	44	46
Relapse rate[b]			
Men			
3-year	3.6	2.6	2.3
6-year	1.9	3.4	2.2
Women			
3-year	3.7	3.3	3.5
6-year	2.6	3.2	3.0

[a]Cessation rate = quitters/baseline smokers.
[b]Relapse rate = relapsers/baseline former smokers.

TABLE 2—Three-Year Smoking Cessation Rates and Adjusted Relative Odds of Cessation According to Selected Characteristics

	Smokers, No.	Cessation Rate, %	Odds Ratio[a] of Quitting	95% CI
Age at baseline, y				
65–69	618	23	1.0	...
70–74	381	30	1.5	1.1, 2.0
75–84	239	36	1.9	1.3, 2.8
≥ 85	21	52	3.5	1.2, 10.0
Race				
Black	86	22	0.7	0.4, 1.2
Non-Black	1173	29	1.0	...
Marital status				
Ever married	1159	28	1.0	...
Never married	98	36	1.6	1.0, 2.5
Cigarettes smoked, no./d				
<10	228	40	1.0	...
10–19	327	30	0.7	0.5, 1.0
20–24	445	25	0.5	0.4, 0.8
≥ 25	224	17	0.3	0.2, 0.6
Conditions reported at baseline[b]				
No	935	29	1.0	...
Yes	324	25	0.7	0.5, 1.0
Interval diagnosis of myocardial infarction, stroke, or cancer				
No	1060	25	1.0	...
Yes	199	44	2.4	1.7, 3.4

[a]Odds ratios were derived from logistic regression models including all main effects listed plus gender, education, years of smoking, alcohol consumption, and site.
[b]Myocardial infarction, stroke, cancer, angina, intermittent claudication.

Discussion

Older adults in these community-based samples had annual smoking cessation and relapse rates of 10% and less than 1%, respectively. Interval diagnosis of myocardial infarction, stroke, or cancer was associated with smoking cessation but not relapse. The oldest old were more likely to quit smoking and less likely to relapse than those 65 to 69 years of age. The association of age and intensity of smoking with quitting and relapse is similar to that in younger populations,[7,9,12] as is lower cessation among Blacks.[12] Our sample included an adequate number of smokers through age 84 but few at older ages as a result of lower smoking prevalence[14] and high mortality.[1,4,19]

Several limitations of the data merit consideration. We lacked data on the number of quit attempts, the method of

TABLE 3—Three-Year Smoking Cessation Rates and Adjusted Relative Odds of Cessation According to Selected Characteristics

	Smokers, No.	Cessation Rate, %	Odds Ratio[a] of Quitting	95% CI
Age at baseline, y				
65–69	492	44	1.0	...
70–74	285	45	0.9	0.7, 1.3
75–84	159	53	1.1	0.7, 1.6
≥ 85	8	88	6.4	0.7, 60.5
Race				
Black	60	33	0.5	0.2, 0.9
Non-Black	884	47	1.0	...
Cigarettes smoked, no./d				
<10	167	58	1.0	...
10–19	257	49	0.6	0.4, 1.0
20–24	335	44	0.5	0.3, 0.8
≥ 25	156	33	0.3	0.2, 0.5
Conditions reported at baseline[b]				
No	721	46	1.0	...
Yes	223	45	0.8	0.6, 1.1
Diagnosis of myocardial infarction, stroke, or cancer between baseline and year 3				
No	836	44	1.0	...
Yes	108	60	1.7	1.1, 2.7
Diagnosis of myocardial infarction, stroke, or cancer between years 3 and 6				
No	782	43	1.0	...
Yes	162	61	2.3	1.6, 3.4

[a]Odds ratios were derived from logistic regression models including all main effects listed plus gender, education, marital status, years of smoking, alcohol consumption, and site.
[b]Myocardial infarction, stroke, cancer, angina, intermittent claudication.

TABLE 4—Three-Year Relapse Rates and Adjusted Relative Odds of Relapse According to Selected Characteristics

	Former Smokers, No.	Relapse Rate, %	Odds Ratio[a] of Relapse	95% CI
Age at baseline, y				
65–69	739	5.1	1.0	...
70–74	619	2.8	0.4	0.2, 0.7
75–84	588	1.9	0.2	0.1, 0.4
≥ 85	104	1.9	0.2	0.04, 0.9
Race				
Black	90	3.3	1.3	0.3, 5.1
Non-Black	1913	3.3	1.0	...
Cigarettes smoked per day when smoking, no.				
<10	370	1.4	1.0	...
10–19	302	3.3	2.2	0.7, 6.7
20–24	550	4.9	3.3	1.2, 9.2
≥ 25	702	3.3	1.9	0.7, 5.4
Age when last smoked regularly, y				
<65	1327	1.4	1.0	...
≥ 65	550	8.4	8.4	4.0, 17.4
Conditions reported at baseline[b]				
No	1420	2.9	1.0	...
Yes	630	4.3	1.3	0.8, 2.3
Diagnosis of myocardial infarction, stroke, or cancer between baseline and year 3				
No	1745	3.4	1.0	...
Yes	305	3.0	0.9	0.4, 1.9

[a]Odds ratios were derived from logistic regression models including all main effects listed plus gender, education, marital status, duration of smoking, alcohol consumption, and site.
[b]Myocardial infarction, stroke, cancer, angina, intermittent claudication.

cessation, and the validity of self-reported smoking status. Self-reported smoking status is usually valid, and many adults successfully quit on their own.[5,10,20] The dates of quitting and disease diagnosis are unknown in some cases. Dynamic patterns were not accounted for because we had limited information about earlier nonsmoking intervals and the length of cessation afterward. The use of baseline and follow-up smoking information misclassified a few quitters but did not affect the significant predictors or our conclusions.

Methodological differences limit comparison of this study with prior work. Prior surveys were cross-sectional[2,5,7,8,10,19,21] or defined cessation differently.[7,10–12] Period differences in cessation rates have been reported.[22] Nevertheless, in younger populations only 33% to 40% quit smoking in a decade.[9,12]

We found, as suggested in middle age,[6,7,12–13] that older adults diagnosed with chronic disease have twofold higher odds of smoking cessation. About 25% of the quitters had a diagnosis of myocardial infarction, stroke, or cancer during the follow-up. The majority who quit do so before or without the onset of these smoking-related illnesses, related in part to demographics, smoking intensity, and possibly to unmeasured factors such as attitudes or other smoking-related chronic conditions or symptoms. Prevalent disease in a baseline smoker, which was inversely related to cessation, may be most common among "hard-core" smokers who did not quit despite their illness. Analysis by duration of illness was not possible.

These findings have important clinical implications for health professionals caring for older smokers. While patients may be highly amenable to smoking cessation around the time of diagnosis of smoking-related illnesses such as atherosclerosis or cancer, primary prevention efforts would have a far greater impact. We join the call for improved cessation initiatives by physicians and other health professionals.[23–25] Given the aging of the US population, which includes younger cohorts with higher smoking prevalence,[19,21] concerted attention to smoking cessation is needed at all ages, including the elderly. □

Acknowledgments

This study was supported in part by contracts N01-AG-0-2105, N01-AG-0-2106, and N01-AG-0-2107 from the National Institute on Aging.

Earlier analyses were presented in part at the 117th and 119th annual meetings of the

American Public Health Association, October 1989, Chicago, Ill, and November 1991, Atlanta Ga.

We thank Drs. Pat Colsher and Sherry L. Mills for helpful comments on an earlier version of the manuscript.

References

1. Centers for Disease Control. Smoking-attributable mortality and years of potential life lost—United States, 1988. *MMWR.* 1991;40:62–63, 69–71.
2. Centers for Disease Control. *Reducing the Health Consequences of Smoking: 25 Years of Progress. A Report of the Surgeon General, 1989.* Rockville, Md: Public Health Service, Office on Smoking and Health; 1989. DHHS publication (CDC) 89-8411.
3. Jajich CJ, Ostfeld AM, Freeman DH Jr. Smoking and coronary heart disease mortality in the elderly. *JAMA.* 1984;252:2831–2834.
4. LaCroix AZ, Lang J, Scherr P, et al. Smoking and mortality among older men and women in three communities. *N Engl J Med.* 1991;324:1619–1625.
5. Centers for Disease Control. *The Health Benefits of Smoking Cessation. A report of the Surgeon General, 1990.* Rockville, Md: Public Health Service, Office on Smoking and Health; 1990. DHHS publication no. (CDC) 90-8416.
6. Hermanson B, Omenn GS, Kronmal RA, Gersh BJ. Beneficial six-year outcome of smoking cessation in older men and women with coronary artery disease: results from the CASS Registry. *N Engl J Med.* 1988; 319:1365–1369.
7. Novotny TE, Fiore MC, Hatziandreu E, Giovino G, Mills SL, Pierce J. Trends in smoking by age and sex, United States, 1974–1987: the implications for disease impact. *Prev Med.* 1990;19:552–561.
8. Kabat GC, Wynder EL. Determinants of quitting smoking. *Am J Public Health.* 1987;77:1301–1305.
9. Sorlie PD, Kannel WB. A description of cigarette smoking cessation and resumption in the Framingham study. *Prev Med.* 1990;19:335–345.
10. Fiore MC, Novotny TE, Pierce JP, et al. Methods used to quit smoking in the United States. Do cessation programs help? *JAMA.* 1990;263:2760–2765.
11. Hatziandreu EJ, Pierce JP, Lefkopoulou M, et al. Quitting smoking in the United States in 1986. *JNCI.* 1990;82:1402–1406.
12. McWhorter WP, Boyd GM, Mattson ME. Predictors of quitting smoking: the NHANES I followup experience. *J. Clin Epidemiol.* 1990;43:1399–1405.
13. Croog SH, Richards NP. Health beliefs and smoking patterns in heart patients and their wives: a longitudinal study. *Am J Public Health.* 1977;67:921-930.
14. Colsher PW, Wallace RB, Pomrehn PR, et al. Demographic and health characteristics of elderly smokers: results from Established Populations for the Epidemiologic Studies of the Elderly. *Am J Prev Med.* 1990;6:61–70.
15. Cornoni-Huntley J, Brock DB, Ostfeld AM, Taylor JO, Wallace RB, eds. *Established Populations for Epidemiologic Studies of the Elderly: Resource Data Book.* Bethesda, Md: National Institutes of Health, National Institute on Aging; 1986. NIH publication no. 86-2443.
16. Cornoni-Huntley JC, Foley DJ, White LR, et al. Epidemiology of disability in the oldest old: methodologic issues and preliminary findings. *Milbank Q.* 1985;63:350–376.
17. Rose GA, Blackburn H, Gillum RF, Prineas RJ. *Cardiovascular Survey Methods.* Geneva, Switzerland: World Health Organization; 1982:162–165.
18. Harrell FE Jr. The LOGIST procedure. In: Hastings RP, ed. *SUGI Supplemental Library User's Guide, Version 5 Edition.* Cary, NC: SAS Institute Inc; 1986;269–293.
19. Harris JE. Cigarette smoking among successive birth cohorts of men and women in the United States during 1900–80. *JNCI.* 1983;71:473–479.
20. Pierce J, Giovino G, Hatziandreu E, Shopland D. National age and sex differences in quitting smoking. *J Psychoactive Drugs.* 1989;21:293–298.
21. Warner KE. Effects of the antismoking campaign: an update. *Am J Public Health.* 1989;79:144–151.
22. Garvey AJ, Bossé R, Glynn RJ, Rosner B. Smoking cessation in a prospective study of healthy adult males: effect of age, time period, and amount smoked. *Am J Public Health.* 1983;73:446–450.
23. Husten CG, Manley MW. How to help your patients stop smoking. *Am Fam Physician.* 1990;42:1017–1026.
24. Rice D, White S, Dreis M, et al. Smoking cessation working group. In: Abdellah FG, Moore SR, eds. *Surgeon General's Workshop: Health Promotion and Aging Proceedings.* Washington, DC: US Department of Health and Human Services, Public Health Service; 1988:88–93.
25. Rimer BK, Orleans CT. The family physician's role in helping older smokers quit. *Am Fam Physician.* 1990;42:959–965.

The Tobacco Institute's "It's the Law" Campaign: Has It Halted Illegal Sales of Tobacco to Children?

Joseph R. DiFranza, MD, and Linda J. Brown

ABSTRACT

A survey of tobacco retailers was conducted to evaluate the efficacy of the Tobacco Institute's "It's the Law" program, which has the goal of eliminating the illegal sale of tobacco to minors. Only 4.5% (7) of the 156 retailers surveyed were participating in the "It's the Law" program. Eighty-six percent of the retailers who were participating in the program were willing to illegally sell cigarettes to children, compared with 88% (131/149) of the retailers who were not participating. (*Am J Public Health.* 1992;82:1271–1273)

Introduction

Nationwide, three out of four tobacco retailers illegally sell tobacco to children.[1,2] Consequently, the tobacco industry has been criticized for profiting from the use of a distribution system that results in the illegal sale of $1.5 billion of its products to children each year.[3,4] The easy access to tobacco provided by this distribution system has been identified as an important preventable factor in adolescent tobacco use.[5]

In an attempt to reduce children's access to tobacco, a number of communities have solicited voluntary merchant compliance with laws prohibiting the sale of tobacco to minors.[6–9] Unfortunately, despite

Joseph R. DiFranza is with the Department of Family and Community Medicine and Linda J. Brown is a medical student at the University of Massachusetts Medical School, Worcester, Mass.

Requests for reprints should be sent to Joseph R. DiFranza, MD, 47 Ashby State Rd, Fitchburg, MA 01420.

This paper was submitted to the Journal November 4, 1991, and accepted with revisions April 7, 1992.

Editor's Note. See related editorial by Giovino et al. (p 1203) in this issue.

a heavy investment of resources, illegal sales have continued to be widespread.[7–9]

By contrast, Woodridge, Ill, has shown that 100% compliance can be achieved quickly and inexpensively with a simple program of compliance checks conducted with the assistance of underage buyers.[5] Most remarkably, the decreased access resulting from active enforcement produced a 50% decline in tobacco use, and a 69% decline in daily tobacco use, among junior high students in Woodridge.[5]

Faced with the national proliferation of effective local legislation concerning the sale of tobacco to minors, the Tobacco Institute on December 11, 1990, announced that it would "help curb youth access to cigarettes through aggressive work with the retail community and by supporting new state laws".[10] Without exception, the new state laws supported by the Tobacco Institute serve to preempt and overturn effective local efforts to control the sale of tobacco to children.[11] For example, a tobacco industry–sponsored bill in Massachusetts would have made it illegal for police to conduct Woodridge-style compliance checks.

The Tobacco Institute's "It's the Law" campaign is marketed as an attempt to "help retailers observe and enforce state laws prohibiting tobacco sales to youth" through the distribution of signs, decals, buttons, employee certification brochures, and other literature.[10]

To determine whether tobacco industry self-regulation is a useful alternative to active enforcement of the law, we sought to test the effectiveness of the "It's the Law" campaign 7 months after its launch.

Methods

Subjects

The subjects for this study were 156 tobacco retailers located by the survey team while driving in central Massachusetts. Retailers located in communities with active enforcement of tobacco access laws were excluded from the study. We had no prior knowledge as to whether retailers participated in the "It's the Law" program or whether they sold tobacco to children.

Procedure

To evaluate merchant compliance with the state law, which prohibits the sale of tobacco to minors under 18 years of age and which requires the posting of warning signs to that effect, we had five underage buyers (aged 13 to 16 years, both male and female) make sham purchases. The prospective buyer would ask the sales clerk for cigarettes, but when the cigarettes were placed on the counter and payment was requested the buyer indicated that he or she did not have enough money and left the premises. With cigarette machines, the underage buyer inserted the correct amount of money and, if he or she had not been stopped at that point, pulled the coin return instead of selecting cigarettes. Buyers were truthful about their age if asked, and made no attempt to appear older. Each establishment was inspected to determine whether it had posted Tobacco Institute "It's the Law" signs or the state-mandated signs. The survey was completed over a 3-week period in the summer of 1991.

Results

The areas surveyed were urban and suburban. Most of the outlets were pharmacies, but convenience stores, gas stations, supermarkets, liquor stores, restaurants, hotels, and a pub were also represented. Eight locations sold cigarettes from vending machines.

Only 4.5% (7) of the 156 retailers surveyed were participating in the Tobacco Institute's "It's the Law" program. Six of the seven participating retailers (86%) were willing to illegally sell cigarettes to children; 88% (131/149) of the retailers who were not participating in the program were willing to sell cigarettes to children. State-mandated warning signs were displayed by 61 retailers, of whom 49 (80%) were willing to illegally sell tobacco to children.

Discussion

We found that the Tobacco Institute's "It's the Law" voluntary compliance program has failed to stop the illegal sale of tobacco products to children. Merchant participation in the program is extremely low, and even participating merchants illegally sell tobacco to children. The participating merchants we interviewed had received their "It's the Law" kits unsolicited, through the mail. By February 1992, the Tobacco Institute claimed to have distributed materials to "thousands of retail outlets across the country" (letter from the Tobacco Institute to Massachusetts Governor William Weld, February 4, 1992).

Health officials should not be deceived by the presence of "It's the Law" signs into thinking that tobacco merchants are complying with the law. Since the industry's attempt at self-regulation has been completely ineffective at reducing children's access to tobacco, the active enforcement of tobacco access laws by health officials is a necessity.

Industry critics have questioned the true motivation behind the "It's the Law" campaign.[12,13] "It's the Law" signs depict a hand brandishing a cigarette and the caption "It's the law/You must be 18 to buy tobacco products." The caption is deceptive, in that in the vast majority of states it is not illegal for minors to purchase tobacco; rather, it is illegal for merchants to sell it to them.[1] Misinformed parents may be afraid to complain about a merchant who is selling tobacco to their children if they mistakenly believe that it is their child who has broken the law.[12] Some observers believe these signs are designed to appeal to teens by presenting tobacco as a forbidden fruit.[12]

The Tobacco Institute has stated that it may spend $10 million on the "It's the Law" campaign over a period of 5 years. However, when we consider that over that 5-year period the tobacco companies will receive over $1 billion in illegal profits from the sale of tobacco to children, we must seriously question the sincerity of their concern.[3]

Active enforcement of tobacco access laws through the use of underage buyers has proved to be a very simple and effective method for decreasing tobacco use among children. The enforcement of these laws is the responsibility of the health community and should not be left to the tobacco industry. □

References

1. DiFranza JR, Norwood BD, Garner DW, Tye JB. Legislative efforts to protect children from tobacco. *JAMA.* 1987;257:3387–3389.
2. Tye JB. Results of trials by minors purchasing cigarettes. *Tobacco Youth Reporter.* 1989;4(2):5.
3. DiFranza JR, Tye JB. Who profits from tobacco sales to children? *JAMA.* 1990; 263:2784–2787.
4. DiFranza JR, Richards JW, Paulman PM, et al. RJR Nabisco's cartoon camel promotes Camel cigarettes to children. *JAMA.* 1991;266:3149–3153.
5. Jason LA, Ji PY, Anes MD, Birkhead SH. Active enforcement of cigarette control laws in the prevention of cigarette sales to minors. *JAMA.* 1991;266:3159–3161.
6. Altman DG. *Tobacco-Free Youth Project: user's Guide for a Program to Reduce the Illegal Sale of Tobacco to Minors.* Palo Alto, Calif: Health Promotion Resources Center, Stanford University, and Stop Teenage Addiction to Tobacco (STAT); 1989.

7. Altman DG, Foster V, Rasenick-Douss L, Tye JB. Reducing the illegal sale of cigarettes to minors. *JAMA.* 1989;261:80–83.
8. Altman DG, Rasenick-Douss L, Foster V, Tye JB. Sustained effects of an educational program to reduce sales of cigarettes to minors. *Am J Public Health.* 1991;81:891–893.
9. Skretny MT, Cummings KM, Sciandra R, Marshall J. An intervention to reduce the sale of cigarettes to minors. *NY State J Med.* 1990;90:54–55.
10. Tobacco Institute. Major new initiative to discourage youth smoking announced. Press Release. Washington, DC; December 11, 1990.
11. DiFranza JR. Untitled article. *Tobacco Access Law News.* 1990;2:1–2.
12. Godshall WT. Blaming victimized children. *SmokeFree Pennsylvania Report.* 1991;1(1):20–21.
13. DuMelle F. Industry drive to cut youths' smoking is doublespeak. *Washington Times.* July 23, 1991.

PHLIS: An Electronic System for Reporting Public Health Data from Remote Sites

Nancy Hargrett Bean, PhD, Stanley Morris Martin, MS, and Henry Bradford, Jr, PhD

ABSTRACT

Disease surveillance conducted by the Centers for Disease Control (CDC) in conjunction with state health departments provides databases of information to public health workers. These databases' utility is limited by the lag time from occurrence of disease events until records are available for analysis. We developed the Public Health Laboratory Information System (PHLIS), a PC-based electronic reporting system for entering, editing, and analyzing data locally and for transmitting data electronically to other state or federal offices. Advantages of PHLIS include reduction in paper handling, decrease in lag time between disease incident and availability of information for analysis, ability to rapidly examine data for clusters of disease, downloadable summary tables, data editing at site of input, data analysis capability, increased interaction among participants, and current data for responses to inquiries. PHLIS is available without cost and is transportable to other agencies, states, or countries. (*Am J Public Health.* 1992;82: 1273–1276)

Introduction

Disease surveillance conducted by the Centers for Disease Control (CDC) with state health departments provides databases that record disease information and data for special studies. These databases provide information to public health workers throughout the United States and other countries. While demand for data is great, utility is limited by the lag time from disease onset until the record is available for analysis. A personal computer–based reporting system was needed for local, county, or state organizations to enter, edit, and analyze data at their own sites and transmit data electronically to other state and federal offices. Although systems exist to address the needs of a single program, a broader system that could be installed in any site (including international), incorporate the needs of divergent organizations, and handle multiple data types was essential. Therefore, we developed the Public Health Laboratory Information System (PHLIS), a flexible system that captures any type of information (e.g., epidemiologic, laboratory, special studies, surveys) from multiple sources (e.g., hospitals, laboratories, state or county offices) and accommodates data transmission regardless of dissimilarities between computer hardware and software at sending and receiving sites.

Several innovative features were designed into PHLIS, for example, a modular approach. Each CDC program responsible for data collection designs and controls each module (a data set containing records concerning a specific pathogen, survey, special study, etc.) in conjunction with state programs. Modules are independent; each can have a unique design (e.g., data entry screen, frequency of data transmission, destination hardware to receive data [personal computer, mainframe, local network, etc.], additional module functions, and analysis and dispersal of data). A decision to change a module does not affect other modules and is limited only by available resources within a program.

Of these features, control and analysis of data are critical. The input of expertise at CDC is maximized by having all programs evaluate and disseminate their data to constituents. This is in contrast to systems that collect, maintain, and evaluate data from multiple sources in one central location. Since CDC collects diverse data for many purposes, it is difficult for any single organization to maintain expertise in all areas.

Nancy Hargrett Bean and Stanley Morris Martin are with the Biostatistics and Information Management Branch, Division of Bacterial and Mycotic Diseases, Center for Infectious Diseases, Atlanta, Ga. Henry Bradford, Jr, is with the Louisiana Public Health Laboratory, New Orleans.

Requests for reprints should be sent to Nancy Hargrett Bean, PhD, Chief, Surveillance & Epidemic Investigation Section, Division of Bacterial and Mycotic Diseases, Centers for Disease Control, CO9, 1600 Clifton Rd, Atlanta, GA 30333.

This paper was submitted to the Journal August 2, 1991, and accepted with revisions March 10, 1992.

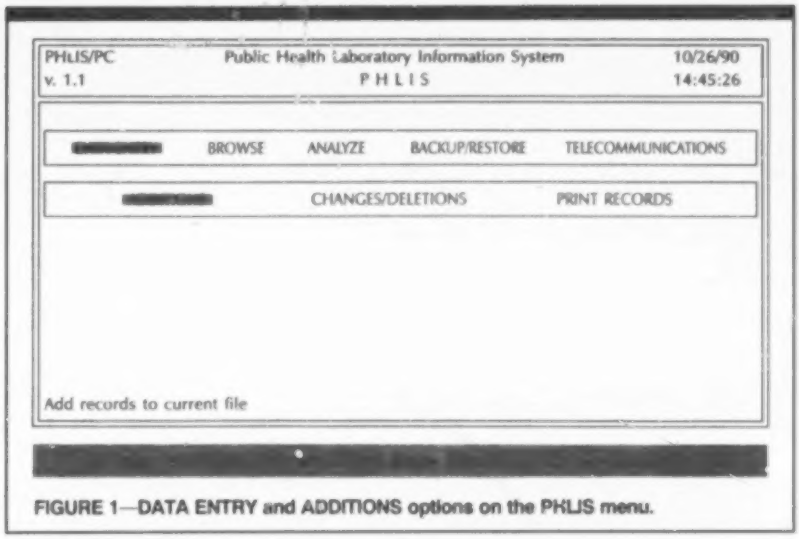


FIGURE 1—DATA ENTRY and ADDITIONS options on the PHLIS menu.

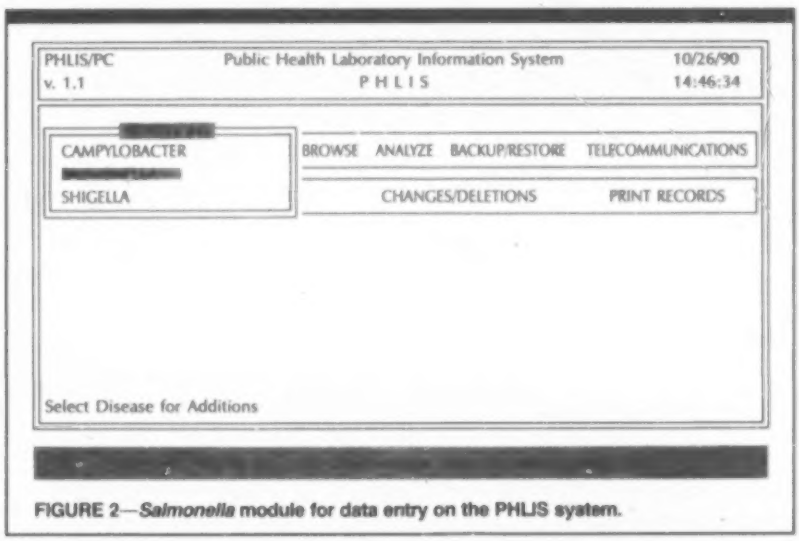


FIGURE 2—*Salmonella* module for data entry on the PHLIS system.

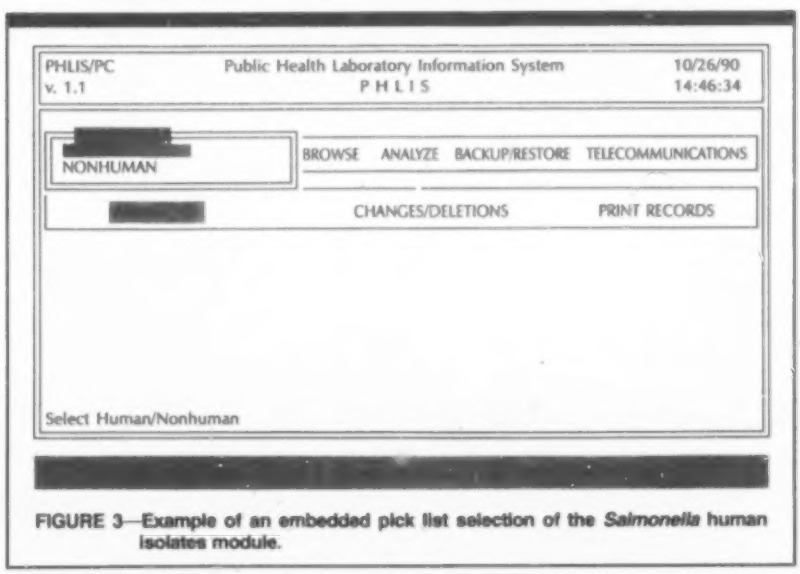


FIGURE 3—Example of an embedded pick list selection of the *Salmonella* human isolates module.

Methods

The Division of Bacterial and Mycotic Diseases, CDC, the Association of State and Territorial Public Health Laboratory Directors, and medical epidemiologists determined which features should be included to meet the needs of potential users. Seven states (Louisiana, New Mexico, Connecticut, Tennessee, Colorado, Delaware, and Maryland) implemented PHLIS using the *Salmonella* surveillance[1] system as the pilot module. Pilot state officials installed and implemented PHLIS with only telephone consultations; no visits to sites by CDC were necessary, and state officials did not travel to CDC for training.

Results

Designed for easy installation and operation, PHLIS requires only a personal computer with DOS[2] 3.0 or higher, a hard disk, at least 512 kilobytes of memory, and a modem. Software for menus provides explicit prompts throughout the system, making PHLIS completely menu driven. PROCOMM PLUS[3] telecommunications software was purchased for participating states. The programs were written using Clipper[4] and Epi Info.[5]

To enter data onto PHLIS select DATA ENTRY, ADDITIONS (Figure 1), and the module (e.g., SALMONELLA). Multiple embedded pick lists are available (e.g., after selection of SALMONELLA [Figure 2], an additional list offers human and nonhuman options [Figure 3]). If *Salmonella* human isolates are selected, the data entry screen for *Salmonella* human isolates appears (Figure 4).

During data entry, the system inhibits input when an error occurs. For example, the serotype is checked for spelling against a list of 1780 serotypes provided in the *Salmonella* module. Other edit features such as range checks and conditional jumps are included.

As a new record is entered, PHLIS assigns a permanent identification number and writes the record to a master file. Transaction files are records not yet sent to other sites, including changes, deletions, and additions. PHLIS recognizes which records have been sent to other sites so that users are not burdened with record tracking. PHLIS adds FIPS county and state codes onto each record as it is entered onto the system.

To change a record, select DATA ENTRY, CHANGES/DELETIONS, a module, and a descriptor (e.g., name). The selected record appears and the user

makes corrections, which are applied to the master file. Corrections and deletions are sent to all receiving sites, including CDC. The BROWSE option provides a way to view but not modify data.

Analysis functions are provided by selecting ANALYZE (Figure 5), a procedure (e.g., LIST VARIABLES), and a module. A listing of module-specific variable names is displayed (Figure 6), eliminating the need to remember variable names. Users are prompted at all steps, enabling novices to perform simple analyses. All Epi Info analysis options are available by selecting OTHER EPI ANALYSIS.

When BACKUP is selected, PHLIS calculates the number of diskettes necessary and allows the user to format additional diskettes. Archived data can be restored by selecting RESTORE, choosing a module, and inserting the archive diskette.

A pyramid structure of data transmission can be developed within states. For example, data initiating locally could be sent to a county or state office and finally to CDC. Alternatively, data can be simultaneously transmitted to state offices and CDC. States retain control over established procedures for data flow, evaluation, and release. The number of participants in a state is not limited, and multiple sites can upload data concurrently.

PHLIS provides several formats for sending data to other sites by specifying the file type (Epi Info, dBASE,[6] ASCII) to be written. Furthermore, if states currently have other types of data systems (e.g., Paradox or VAX), files (ASCII, Epi Info, or dBASE) may be imported into PHLIS in a prescribed format. For example, PHLIS is currently receiving data from two states with VAX systems. Therefore, PHLIS data can be accessed by other software and can incorporate data from other systems.

Data are received via a personal computer dedicated to communications. Selecting SEND TO CDC (Figure 7) and a module causes PHLIS to call the communications computer, log the user on with appropriate identification, upload the data, indicate whether data have been transferred correctly, and log off. No experience is needed with telecommunications to accomplish transfer of data electronically. Data may be transmitted as often as necessary (multiple times in the same day if desired).

Transmitted data are downloaded to the CDC mainframe, where processing and edits are performed by the responsible program. If errors are detected, edits are transmitted back to reporting sites. Messages and tables summarizing data received at CDC are downloaded to sites. The TELECOMMUNICATIONS option (Figure 8) provides access to other CDC systems, and bulletin board script files contain telephone numbers that PHLIS accesses when connecting to other systems. Therefore, PHLIS can connect to additional services as requested by state officials (e.g., E-Mail).

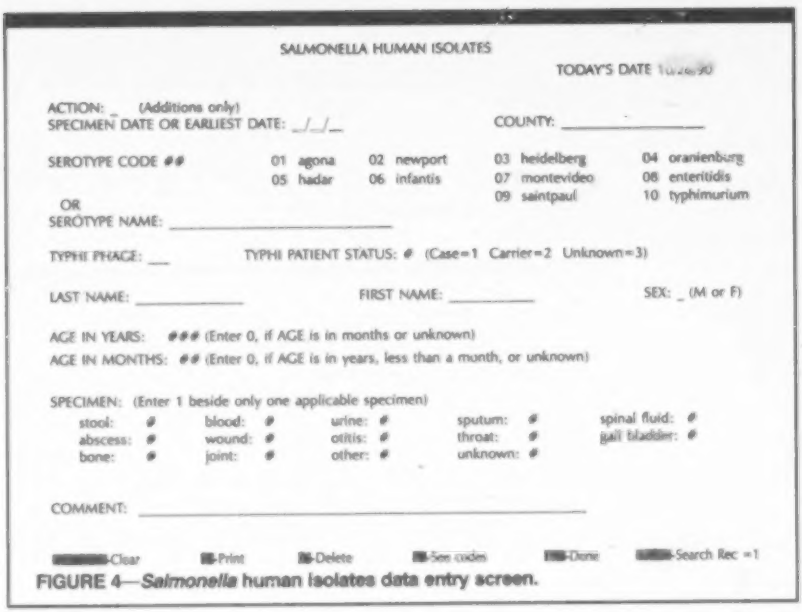

SALMONELLA HUMAN ISOLATES
TODAY'S DATE
ACTION: _ (Additions only)
SPECIMEN DATE OR EARLIEST DATE: _/_/_
COUNTY:
SEROTYPE CODE ##
01 agona 02 newport 03 heidelberg 04 oranienburg
05 hadar 06 infantis 07 montevideo 08 enteritidis
09 saintpaul 10 typhimurium
OR
SEROTYPE NAME:
TYPHI PHAGE:
TYPHI PATIENT STATUS: # (Case=1 Carrier=2 Unknown=3)
LAST NAME:
FIRST NAME:
SEX: _ (M or F)
AGE IN YEARS: ### (Enter 0, if AGE is in months or unknown)
AGE IN MONTHS: ## (Enter 0, if AGE is in years, less than a month, or unknown)
SPECIMEN: (Enter 1 beside only one applicable specimen)
stool: # blood: # urine: # sputum: # spinal fluid: #
abscess: # wound: # otitis: # throat: # gall bladder: #
bone: # joint: # other: # unknown: #
COMMENT:
Clear
Print
Delete
See codes
Done
Search Rec =1

FIGURE 4—*Salmonella* human isolates data entry screen.

Discussion

PHLIS has immediate benefits for public health: (1) Electronic transmission of reports reduces the paper and data entry burdens within states (in one state, the same reports of *Salmonella* are routinely entered into three computers and again at CDC). (2) Access to timely national summaries ensures that data are current and responses to inquires at local, state, and CDC levels are timely. (3) PHLIS increases interaction among public health officials, stimulates more thorough reporting, and fosters greater awareness of reporting problems. Finally, rapid reporting provides the potential to quickly identify unusual disease clusters and alert state epidemiologists and CDC personnel to outbreak episodes. However, the need for careful field investigation and data assess-

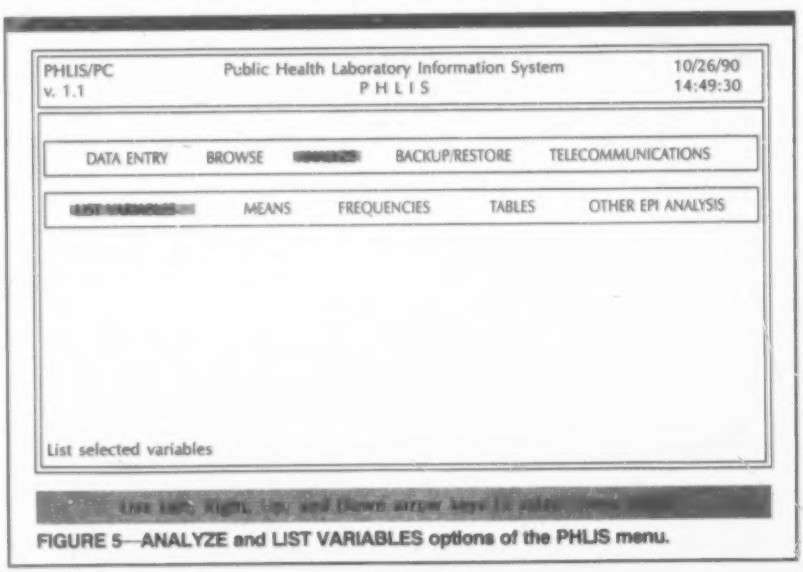
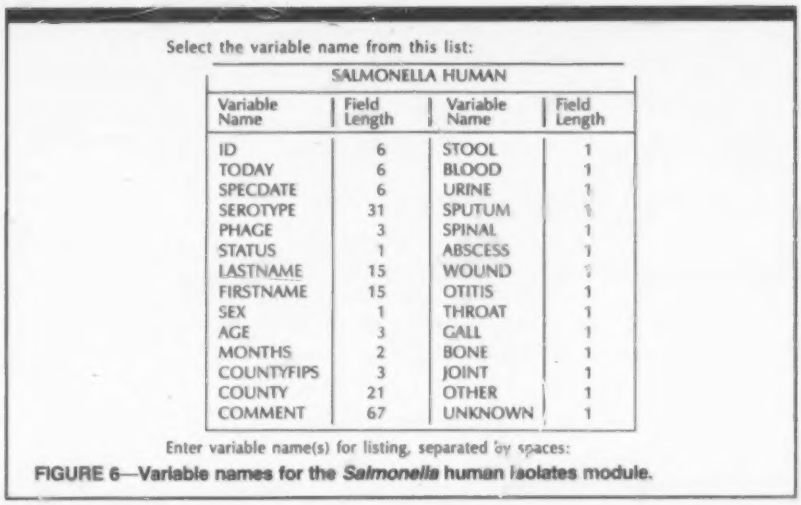


FIGURE 5—ANALYZE and LIST VARIABLES options of the PHLIS menu.



Variable Name	Field Length	Variable Name	Field Length
ID	6	STOOL	1
TODAY	6	BLOOD	1
SPECDATE	6	URINE	1
SEROTYPE	31	SPUTUM	1
PHAGE	3	SPINAL	1
STATUS	1	ABSCESS	1
LASTNAME	15	WOUND	1
FIRSTNAME	15	OTITIS	1
SEX	1	THROAT	1
AGE	3	GALL	1
MONTHS	2	BONE	1
COUNTYFIPS	3	JOINT	1
COUNTY	21	OTHER	1
COMMENT	67	UNKNOWN	1

FIGURE 6—Variable names for the *Salmonella* human isolates module.

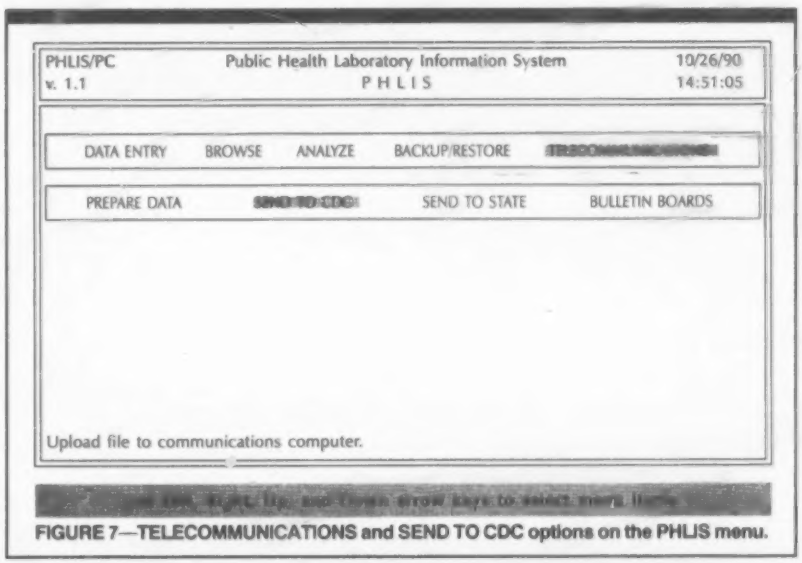


FIGURE 7—TELECOMMUNICATIONS and SEND TO CDC options on the PHLIS menu.

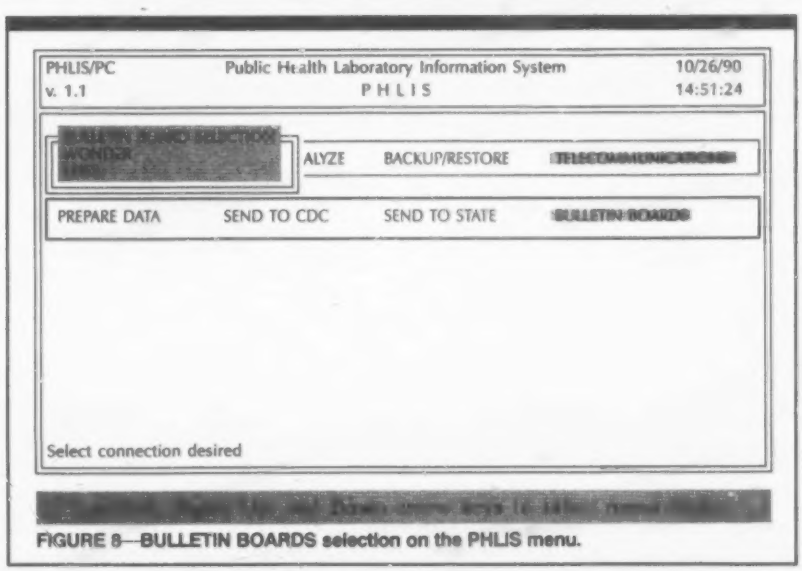


FIGURE 8—BULLETIN BOARDS selection on the PHLIS menu.

ment prior to data entry remains the responsibility of the originating site.

PHLIS can be implemented without the telecommunications components; data can be transferred as ASCII files via modems among sites without a central computer to receive data; and organizations without modems can transmit data via diskettes without affecting system integrity.

PHLIS is currently installed in 41 of the 45 state laboratories that have requested the system. Although the initial PHLIS module was implemented using laboratory-based data, the system accommodates all data types (e.g., epidemiological data). Three additional laboratory modules have been added to the system (*Shigella, Campylobacter,* and *Mycobacteria*), as well as one module to collect epidemiological data from state epidemiologists' offices (seafood case-control study).

PHLIS was designed to provide flexibility with few constraints placed on module design and data transfer. Plans for enhancements include development of an "applications generator" that allows users to develop new modules or make modifications to accommodate unique site needs. PHLIS is available without cost to any organization wishing to implement the system. □

Acknowledgments

Portions of this material were presented at the 32nd Annual Planning Conference of the Association of State and Public Health Laboratory Directors, 1989, and the annual meeting of the Council of State and Territorial Epidemiologists, 1990.

We thank the Management Information Systems Committee of the Association of State and Territorial Public Health Laboratory Directors for their guidance and support, Sara Lyle for computer programming for PHLIS, and the personnel in the seven pilot states for their participation.

References

1. Salmonella *Annual Summary.* Atlanta, Ga: Centers for Disease Control; 1982.
2. IBM Corporation and Microsoft Inc. *Disk Operating System Version 3.0.* IBM Corporation and Microsoft Inc; 1984.
3. Datastorm Technologies Inc. *PROCOMM PLUS Intuitive Communications.* Columbia, Mo: Datastorm Technologies Inc; 1988.
4. Nantucket Corporation. *Clipper, Version Summer '87.* Los Angeles, Calif: Nantucket Corporation; 1987.
5. Dean AD, Dean JA, Burton JH, Dicker RC. *Epi Info, Version 3: A Word Processing, Database, and Statistics Program for Epidemiology for Microcomputers.* Atlanta, Ga: Centers for Disease Control; 1988.
6. Ashton-Tate. *dBASE III Plus.* Torrance, Calif: Ashton-Tate; 1986.

Homosexual and Bisexual Men's Perceptions of Discrimination in Health Services

Nancy E. Kass, ScD, Ruth R. Faden, PhD, MPH, Robin Fox, MS, and Jan Dudley, MPH

ABSTRACT

Questionnaires were distributed to homosexual and bisexual male participants in the Multicenter AIDS Cohort Study and to homosexual and bisexual male patients with acquired immunodeficiency syndrome (AIDS) to determine whether the men believed they had been refused medical or dental treatment because of their sexual orientation or a condition related to the human immunodeficiency virus (HIV). Men with AIDS were significantly more likely (18%) to report being refused treatment by a doctor or dentist on the basis of a known or suspected HIV-related condition than were men who were seropositive (5%) or seronegative (1%). Significantly more respondents reported refusal of dental care than of medical care. (*Am J Public Health.* 1992;82:1277–1279)

Introduction

Since the beginning of the human immunodeficiency virus (HIV) epidemic, there have been reports of discrimination toward HIV-infected persons in many settings,[1–5] including the health care setting.[6–8] Moreover, there have been reports of discrimination in the provision of health services on the basis of patients' sexual orientation; this discrimination is thought to be at least partially related to the association between male homosexuality and risk of HIV infection. Refusal to provide health services is an important ethical issue and is contrary to professional standards. Moreover, refusal of treatment becomes a public health issue when unjustified discrimination results in diminished access to health services.

Although various studies have documented physicians' and dentists' attitudes toward HIV-infected patients,[9–12] and homosexual patients,[9,13] little empirical work has been conducted to determine the experiences of at-risk individuals seeking medical or dental care. It was the purpose of this study to document the extent to which homosexual and bisexual men believe they have been refused services by a doctor or dentist because of their sexual orientation or because of a known or suspected HIV-related condition.

Methods

Study Populations

Three different samples were used in this study. Samples 1 and 2 were recruited from the Baltimore and Los Angeles sites, respectively, of the Multicenter AIDS Cohort Study (MACS). Sample 3 was recruited from gay and bisexual male AIDS patients at The Johns Hopkins Hospital in Baltimore.

The MACS, begun in 1984, is a prospective study of the natural history of HIV infection in gay and bisexual men. Criteria for initial inclusion in the MACS were being male, having a history of homosexual or bisexual activity, and not having acquired immunodeficiency syndrome (AIDS) at the time of enrollment. Most of the men in this volunteer cohort have college degrees (Baltimore, 67%; Los Angeles, 55%) and are White (Baltimore, 92%; Los Angeles, 91%).[14]

MACS appointments are scheduled every 6 months. Baltimore and Los Angeles MACS participants were asked to complete the questionnaire for this study during their eighth visit (between October 1987 and July 1988). In Baltimore, 857 men attended the clinic during the study period. There was a response rate of 93%; therefore, 797 participants constituted sample 1. Because other studies were being conducted at the Los Angeles site during that time, there was only a 19% response rate on site. Therefore, questionnaires were mailed to the remaining Los Angeles MACS participants. There was an overall response rate of 59% from the 1250 participants scheduled for a visit during this period. Thus, sample 2 consisted of 708 Los Angeles MACS participants.

Sample 3 consisted of gay and bisexual male inpatients and outpatients with AIDS at The Johns Hopkins Hospital; these patients were recruited because the number of participants in the MACS with AIDS was small. Patients who were too sick to complete a questionnaire were interviewed. Of 63 patients approached, 59 (94%) agreed to participate, constituting sample 3.

Nancy E. Kass and Ruth R. Faden are with the Program in Law, Ethics, and Health and Robin Fox is with the Department of Epidemiology, all at The Johns Hopkins School of Hygiene and Public Health, Baltimore, Md. Ruth R. Faden is also with the Kennedy Institute of Ethics, Georgetown University, Washington, DC. Jan Dudley is with the Los Angeles Multicenter AIDS Cohort Study, UCLA School of Public Health, Los Angeles, Calif.

Requests for reprints should be sent to Nancy E. Kass, ScD, Program in Law, Ethics, and Health, 624 N Broadway, Baltimore, MD 21205.

This paper was submitted to the Journal August 21, 1991, and accepted with revisions March 20, 1992.

TABLE 1—Percentages of Homosexual and Bisexual Men Reporting Treatment Refusals by Doctors and Dentists, by Antibody Status

	HIV− (n = 852)	HIV+ (n = 585)	PWAs (n = 96)
Has a doctor ever refused to treat you because s/he knew or thought that you were gay or bisexual?*	2	4	3
Has a doctor ever refused to treat you because s/he suspected that you were HIV positive or had an AIDS-related illness?**	1	3	5
Has a dentist ever refused to treat you because s/he knew or thought that you were gay or bisexual?***	1	3	4
Has a dentist ever refused to treat you because s/he suspected that you were HIV positive or had an AIDS-related illness?****	1	4	16

*$P < .05$ for difference between seronegative men (SNs) and seropositive men (SPs) (odds ratio [OR] = 2.5; 95% CI = 1.3, 4.8).
**$P < .01$ for difference between persons with AIDS (PWAs) and SNs (OR = 7.8; 95% CI = 2.9, 26.1); $P < .01$ for difference between SPs and SNs (OR = 3.9; 95% CI = 1.5, 10.1).
***$P < .05$ for difference between PWAs and SNs (OR = 3.7; 95% CI = 1.1, 11.9); $P < .05$ for difference between SPs and SNs (OR = 2.4; 95% CI = 1.1, 5.2).
****$P < .001$ for difference between PWAs and SNs (OR = 22.2; 95% CI = 8.8, 55.9); $P < .001$ for difference between SPs and SNs (OR = 4.7; 95% CI = 2.0, 11.0).

The Research Instrument

Questions for this investigation constituted one section of a larger self-administered questionnaire concerning insurance status and perceived discrimination. In this section, participants were asked four yes/no questions concerning treatment refusal by doctors or dentists (Table 1).

Data Obtained from MACS Records

Data collected specifically for this study were linked with data collected by the MACS at the same or most recent clinic visit, including demographic characteristics; HIV antibody status; AIDS diagnosis; CD4 count; and self-reported weight loss, fevers, fatigue, and diarrhea. (Because the men in sample 3 were not part of the MACS, questionnaires distributed to this group included questions on demographics.)

Statistical Analyses

Two-tailed Fisher's Exact Tests were used to compare respondents with nonrespondents with respect to demographic and health-related characteristics and to test the association between each type of treatment refusal and (1) HIV antibody status and (2) certain demographic characteristics. For bivariate relationships found to be associated with a probability higher than 95%, odds ratios and 95% confidence intervals were calculated. Logistic regression was used to develop predictive models.

Results

Comparison of Respondents with Nonrespondents

Respondents and nonrespondents from Samples 1 and 2 were compared using available MACS data. Respondents were more likely to be White, well educated, and older. Nonrespondents were more likely to be seropositive or to have AIDS. There were no significant differences between respondents and nonrespondents in terms of selected symptoms, CD4 count, or history of hospitalization since their last visit.

Comparison of MACS Participants with Johns Hopkins Participants

Persons with AIDS (PWAs) from the MACS were compared with PWAs recruited from The Johns Hopkins Hospital in terms of educational background and employment status. Although no differences were significant at $P < .05$, PWAs recruited from Johns Hopkins tended to be somewhat less well educated (67% had at least some college education, compared with 88% of PWAs from the MACS) and were more likely to be employed full-time (78%, compared with 67% of PWAs from the MACS).

Treatment Refusals Based on Sexual Orientation

Two percent of the seronegative men (SNs) reported having been refused treatment by a doctor because of the respondent's sexual orientation, compared with 4% of seropositive men (SPs) who did not have AIDS ($P < .05$) and 3% of PWAs. One percent of SNs reported having been refused treatment by a dentist because of the respondent's sexual orientation, compared with 3% of SPs ($P < .05$) and 4% of PWAs ($P < .05$) (Table 1).

Treatment Refusals Based on HIV-Related Conditions

One percent of SNs reported having been refused treatment by a doctor because of a known or suspected HIV-related condition, compared with 3% of SPs ($P < .01$) and 5% of PWAs ($P < .01$). One percent of SNs reported having been refused treatment by a dentist because of a known or suspected HIV-related condition, compared with 4% of SPs ($P < .001$) and 16% of PWAs ($P < .001$).

One percent of SNs reported having been refused treatment by either a doctor or a dentist because of a known or suspected HIV-related condition, compared with 5% of SPs ($P < .001$) and 18% of PWAs ($P < .001$). The same degree of treatment refusal was reported by PWAs from the MACS and by PWAs from The Johns Hopkins Hospital. In the logistic model, having AIDS and being seropositive each contributed independently to predicting treatment refusal from a doctor or dentist on the basis of a known or suspected HIV-related condition; age, education, and symptoms were not significant predictors.

Discussion

Whereas homosexual and bisexual men without AIDS very rarely reported treatment refusals, 18% of PWAs reported a refusal from a doctor or dentist on the basis of a known or suspected HIV-related condition. However, refusal of dental care was considerably more likely to be reported than refusal of medical care.

A limitation of this study is that all data are self-reported, and responses were not validated. We do not know the circumstances of the incidents, or if the practitioners would have described the interactions in the same way as did the respondents. Participants may have characterized any difficulty they had in securing access to services as a refusal when other reasons truly existed; moreover, men who had ex-

perienced refusal may have been more likely to respond to our questionnaire. Either of these possibilities could result in overreporting. However, many participants volunteered that they had deliberately sought care from gay practitioners to avoid discrimination problems. In addition, SPs and PWAs were more heavily represented among nonrespondents in our sample. Both of these factors could lead to underreporting of treatment refusal.

Treatment refusals may exacerbate the existing disproportionate public burden of care and treatment costs for persons infected with HIV. The disproportionate number of PWAs in this study who reported a treatment refusal, particularly by dentists, is cause for concern and warrants public health and public policy attention. □

Acknowledgments

This study was supported by grant HS 06108 from the National Center for Health Services Research and Health Care Technology Assessment and grants AI-72634, AI-726311, and 5 M01-RR-00722 from the National Institutes of Health.

Data from this paper were presented at the Fifth International Meeting on AIDS, San Francisco, Calif, June 20–24, 1990.

References

1. Gostin L. The AIDS Litigation Project: a national review of court and human rights commission decisions, part II: discrimination. *JAMA.* 1990;263:2086–2093.
2. Blendon RJ, Donelan K. Discrimination against people with AIDS. *N Engl J Med.* 1988;319:1022–1026.
3. Robertson v Granite City Comm Unit Sch Dist No. 9, 684 F Supp 1002 (SD Ill 1988).
4. NY v Dow, USDC Conn. *AIDS Lit Rpt* 6/10/88.
5. Ritchie v Colonial Cheshire II. *The Lambda Update.* Spring 1991.
6. Elstein v NYS Division of Human Rights, Superior Court, Onondaga County, NY, February 14, 1989.
7. Doe v Shasta Hosp. *The Lambda Update.* Summer 1990.
8. Hurwitz v NYC Commission on Human Rights, 142 Misc 2d 214, 535 NYS 2d 1007 (1988).
9. Hayward RA, Shapiro MF. A national study of AIDS and residency training: experiences, concerns, and consequences. *Ann Intern Med.* 1991;114:23–32.
10. Kelly JA, St. Lawrence JS, Smith S, Hood HV, Cook DJ. Stigmatization of AIDS patients by physicians. *Am J Public Health.* 1987;77:789–791.
11. Shin DM, Wong AY. AIDS/HIV: infection control practices, attitudes and beliefs of dental faculty and students. Presented at the Fifth International Conference on AIDS; June 20–24, 1990; San Francisco, Calif.
12. Levin BW, Driscoll JM, Fleischman AR. Treatment choice for infants in the neonatal intensive care unit at risk for AIDS. *JAMA.* 1991;265:2976–2981.
13. Kelly JA, St. Lawrence JS, Smith S, Hood HV, Cook DJ. Medical students' attitudes toward AIDS and homosexual patients. *J Med Educ.* 1987;62:549–556.
14. Kaslow RA, Ostrow DG, Detels R, Phair JP, Polk BF, Rinaldo CR, Jr (for the Multicenter AIDS Cohort Study). The Multicenter AIDS Cohort Study: rationale, organization, and selected characteristics of the participants. *Am J Epidemiol.* 1987; 126:310–318.

Radiologic Society Conference to Be Held in Late November

The 78th Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA) will be held November 27 to December 4, 1992, at McCormick Place in Chicago, Ill. More than 1400 scientific papers will be presented this year, from clinical studies in radiation therapy to new procedures in ultrasound, MRI, CT, SPECT, and PET scans. In addition, there will be nearly 1000 technical and scientific exhibits.

More than 50 000 health care professionals from more than 80 countries are expected to attend, making the conference the largest radiology meeting in the world. For registration, contact RSNA at 2021 Spring Rd, Suite 600, Oak Brook, IL 60521; tel. (708) 571-2670, fax (708) 571-7837.

NOTE

Some charts and graphs were color-coded in the original document. While it has not been possible to reproduce these items in color, their equivalent values on the black-and-white scale are often distinct enough to convey the same information as the original colors.

Congressional Information Service, Inc
4520 East-West Highway
Bethesda, Maryland 20814

CIS

NOTE

ONLY THE FOLLOWING PAGES OF THIS

PUBLICATION HAVE BEEN FILMED

Cov. 1, Pgs: 1313, 1314,

1328-1398

A2623-1

American Journal of Public Health

October 1992, Vol. 82, No. 10 Established 1911

86389
028
ED KIANKA
CONGRESSIONAL INFORMATION SVC
4520 EAST-WEST HIGHWAY
BETHESDA MD 20814

Editorials

1323 Induced Abortion and Health as a Value *M. Susser*

1325 RU 486 *A. Rosenfield*

1326 Global Microbial Traffic and the Interchange of Disease *S. S. Morse*

Public Health Policy Forum

1328 Commentary: The Public Health Consequences of Restricted Induced Abortion—Lessons from Romania *P. Stephenson, M. Wagner, M. Badea, and F. Serbanescu*

Featuring Maternal and Child Health

1332 Determinants of Depressive Symptoms in the Early Weeks after Miscarriage *R. Neugebauer, J. Kline, P. O'Connor, P. Shrout, J. Johnson, A. Skodol, J. Wicks, and M. Susser*

1340 Trends in Obstetric Operative Procedures, 1980 to 1987 *S. C. Zahniser, J. S. Kendrick, A. L. Franks, and A. F. Saftlas*

1345 Changes in Mammography Use: Economic, Need, and Service Factors *J. G. Zapka, D. Hosmer, M. E. Costanza, D. R. Harris, and A. Stoddard*

1352 Current Cigarette Smoking and Risk of Acute Pelvic Inflammatory Disease *D. Scholes, J. R. Daling, and A. S. Stergachis*

1356 Lead Exposure and Child Behavior *W. G. Sciarillo, G. Alexander, and K. P. Farrell*

1361 Pregnancy among the Hmong: Birthweight, Age, and Parity *D. Helsel, D. B. Petitti, and P. Kunstadter*

1365 Neonatal Tetanus in Rural Bangladesh: Risk Factors and Toxoid Efficacy *W. G. Hlady, J. V. Bennett, A. R. Samadi, J. Begum, A. Hafez, A. I. Tarafdar, and J. R. Boring*

1370 Improving Estimates of HIV-1 Seroprevalence among Childbearing Women: Use of Smaller Blood Spots *N. J. Hoxie, J. M. Vergeront, J. R. Pfister, G. L. Hoffman, P. A. Markwardt-Elmer, and J. P. Davis*

1377 Legal Blindness among 10-Year-Old Children in Metropolitan Atlanta: Prevalence, 1985 to 1987 *C. D. Drews, M. Yeargin-Allsopp, C. C. Murphy, and P. Decoufle*

1380 Sudden Infant Death Syndrome and Maternal Smoking *M. H. Malloy, H. J. Hoffman, and D. R. Peterson*

1383 Stage at Diagnosis in Breast Cancer: Race and Socioeconomic Factors *B. L. Wells and J. W. Horm*

1386 Measuring the Use of Mammography: Two Methods Compared *D. Degnan, R. Harris, J. Ranney, D. Quade, J. A. Earp, and J. Gonzalez*

1388 Women with Multiple Sexual Partners: United States, 1988 *S. N. Seidman, W. D. Mosher, and S. O. Aral*

(continued page 1313)

APHA
American Public Health Association

Featuring Maternal and Child Health

Contents

American Journal of Public Health

The Journal's one-page guidelines for authors, "What *AJPH* Authors Should Know," are printed in each issue. They can also be obtained from the Editorial Office.

October 1992, Vol. 82, No. 10 ISSN: 0090-0036 Established 1911

Article

1374 The Disclosure of Celebrity HIV Infection: Its Effects on Public Attitudes *S. C. Kalichman and T. L. Hunter*

Public Health Brief

1394 The Medical Origins of Homelessness *M. A. Winkleby, B. Rockhill, D. Jatulis, and S. P. Fortmann*

Health Law and Ethics

1399 The Search for Meaning: RU 486 and the Law of Abortion *S. S. Banwell and J. M. Paxman*

Public Health Then and Now

1407 The Interchange of Disease and Health between the Old and New Worlds *G. Berlinguer*

Notes from the Field

1414 Controlling Cardiovascular Disease: The Role of the Local Health Department *R. Brownson, C. Smith, N. Jorge, C. Dean, and L. DePrima*

1416 Implementing the Veterans Administration's No-Smoking Policy *C. Bevvino and M. C. Jacob*

Other Journal Departments

1314 APHA Masthead

1318 Advertisers' Index

1322 What *AJPH* Authors Should Know

1323 Journal Masthead

1418 Letters to the Editor

1422 Book Corner

1429 Annual Meeting Registration Form and Information

1433 Job Opportunities

American Public Health Association

American Journal of Public Health

American Public Health Association

The American Journal of Public Health, **published monthly, is the official Journal of the American Public Health Association, 1015 Fifteenth Street, NW, Washington, DC 20005; (202) 789-5600. APHA is a nongovernmental professional society founded in 1872, representing all disciplines and specialties in public health.**

APHA OFFICERS

Joyce C. Lashof, MD, *President*

C. William Keck, MD, MPH, *Immediate Past President*

Helen Rodriguez-Trias, MD, *President-Elect*

Arnold E. Greenberg, SM, *Vice-President (USA)*

Trevor Hancock, MB, MHSc, *Vice-President (Canada)*

Julio Teja Perez, MD, *Vice-President (Latin America)*

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Frances M. Veverka, RS, MPH, *Chair, Executive Board*

George E. Hardy, Jr, MD, MPH, *Speaker, Governing Council*

William H. McBeath, MD, MPH, *Executive Director*

EXECUTIVE BOARD

Frances M. Veverka, RS, MPH (1992) *Chair*

Marilyn Aguirre-Molina, EdD, MS (1993)

Brian A. Dobrow, MPH (1992)

Caswell A. Evans, Jr, DDS, MPH, *Treasurer*

Kristine M. Gebbie, BSN, MN (1994) *Vice-Chair*

Alan R. Hinman, MD, MPH (1995)

Janet Heinrich, DrPH, MPH (1993)

Margaret H. Jordan, RN, MPH (1992)

C. William Keck, MD, MPH, *Immediate Past President*

Margo D. Kerrigan, MPH (1995)

Joyce C. Lashof, MD, *President*

Barry S. Levy, MD, MPH (1993)

Helen Rodriguez-Trias, MD, *President-Elect*

Allan G. Rosenfield, MD (1994)

Reed V. Tuckson, MD (1994)

Steven Uranga McKane, DMD, MPH (1995)

EX OFFICIO

William H. McBeath, MD, MPH, *Executive Director*

Joanne E. Lukomnik, MD, *Chair, Action Board*

George Pickett, MD, MPH, *Chair, Program Development Board*

George E. Hardy, MD, MPH, *Speaker, Governing Council*

The American Journal of Public Health, published monthly, is the official Journal of the American Public Health Association, Inc, 1015 Fifteenth St, NW, Washington, DC 20005. Founded in 1872, the American Public Health Association is organized to protect and promote personal and environmental health by exercising leadership in the development and dissemination of health policy. Representing all disciplines and specialties in public health, APHA is the largest public health association in the world, with an aggregate membership of 50 000, including its state and local affiliates.

Manuscripts, Correspondence, and Letters to the Editor: Send to Mervyn Susser, DPH, Editor, *American Journal of Public Health,* 1015 Fifteenth St, NW, Washington, DC 20005.

Information on Manuscript Preparation is published monthly in the Journal. The Journal publishes reports of original research, research methods, program evaluations, analytic reviews, including health policy analysis, and reports for special departments. A paper is considered with the understanding that it or its essential substance neither has been published nor is under consideration elsewhere.

Subscription price, individual, is $100/year US, $140/year international; **institutional,** $160/year US, $200/year international. Single copies $10 US, $14 international (surface delivery), $22 international (air delivery). Prepaid only.

Postal Information: Second class postage paid at Washington, DC, and additional mailing offices. Postmaster: Send address changes to *American Journal of Public Health,* 1015 Fifteenth St, NW, Washington, DC 20005. AJPH allotted $40; TNH allotted $6 of APHA membership dues. Information on individual membership appears elsewhere in this Journal, or contact Director, Membership Services, at APHA headquarters.

Advertising Sales: Mary Beth Schultz, Advertising Manager, APHA, 1015 Fifteenth St, NW, Washington, DC 20005; (202) 789-5668 or 5600. Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or APHA.

Reprints: Request single reprints of articles published in the Journal from the authors at the address given in the footnote on first page of published article. For cost information on 100 or more reprints, contact Marilyn Butler, AJPH Reprints, APHA, 1015 Fifteenth St, NW, Washington, DC 20005; (202) 789-5649.

Back Issues of the Journal: A limited inventory of back issues and volumes of the journal is available at current subscription rates plus handling charge. All orders under $35 must be prepaid. Vols 1 to present are also available on microfilm.

Copyright © 1992 by the American Public Health Association, Inc. All material subject to this copyright may be photocopied for the noncommercial purposes of scientific or educational advancement. Opinions expressed by authors of articles summarized, quoted, or published in full in this Journal represent the opinions of the authors and do not necessarily reflect the official policy of the American Public Health Association or the institution with which the author(s) is (are) affiliated, unless so specified.

Any report, article, or paper prepared by employees of the US Government as part of their official duties is, under the Copyright Act, a "work of the United States Government" for which copyright protection under Title 17 of the US Code is not available. However, the Journal format is copyrighted and pages may not be photocopied without permission of the American Public Health Association, except in limited quantities for educational purposes, as stated above.

Copying done for other than personal or internal reference use—such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale—without the expressed permission of the American Public Health Association is prohibited. Requests for special permission or bulk orders should be addressed to APHA headquarters.

The American Journal of Public Health is indexed in whole or in part in these data bases: Bibliographic Index of Health Education Periodicals, Biological Abstracts, Biology Digest, BRS/COLLEAGUE, Chemical Abstracts, Cumulative Index to Nursing & Allied Health Literature, Environmental Information Center Inc, Environmental Periodicals Bibliography, Excerpta Medica, Grateful Med, Hospital Literature Index, Hospital Management Review, Human Resources Abstracts, Index Medicus (Abridged), MEDLARS, MEDSOC, Nuclear Science Abstracts, Physical Education Index, Public Affairs Information Service, Public Health Reviews, Safety Science Abstracts, Sage Family Studies Abstracts, Science Citation Index, and Social Work Research and Abstracts.

Public Health Policy Forum

Commentary: The Public Health Consequences of Restricted Induced Abortion—Lessons from Romania

Patricia Stephenson, ScD, Marsden Wagner, MD, MSPH, Mihaela Badea, MD, and Florina Serbanescu, MD

ABSTRACT

The question of whether abortion should be legal is currently being decided in many countries. Although much of the discussion has focused on ethical issues, the public health consequences should not be overlooked and should be addressed realistically and responsibly. Nowhere are the public health manifestations of restricted abortion more apparent than in Romania. The pronatalist policies of the Ceaucescu regime resulted in the highest maternal mortality rate in Europe (approximately 150 maternal deaths per 100 000 live births) and in thousands of unwanted children in institutions. (*Am J Public Health.* 1992;82:1328–1331)

Introduction

Abortion policy is being debated in a number of countries in which abortion is currently legal and women have access to safe abortion, performed by qualified practitioners, through the health care system. Some political or religious groups in Hungary, Poland, Germany, Spain, Sweden, and the United States seek to ban abortion or at least to restrict it as much as possible.

Both proponents and opponents of legal abortion debate the issue in the abstract terms of women's rights, fetal rights, religious theories, or constitutional principles. However, those who ultimately decide this question will sooner or later have to come to terms with the concrete reality of what can happen to women and children when and if the right to safe abortion is taken away.

Safe Abortion and Maternal Mortality

When access to safe abortion has been introduced in a country, maternal mortality has decreased. In the United States during the 1960s, as some states began to change their laws on abortion and hospital abortion policies became less restrictive, gradual decreases in maternal mortality were noted. After the US Supreme Court decision in 1973 guaranteeing women the right to safe abortion, national maternal mortality rates decreased further. In England and Wales there were no deaths due to unsafe abortion in the triennium 1982 to 1984, compared with 75 to 80 such deaths per triennium prior to the Abortion Act of 1967 (which gave free access to safe abortion).[1] Although it is reasonable to conclude that these trends are due to the changes in the abortion laws, one can never completely rule out the possible simultaneous effects of other factors.

More instructive is what happens in a country when abortion is made illegal and access to safe abortion taken away. Before 1966, Romanian women—like their neighbors in other Eastern European countries—had access to safe abortion through the country's health care system. In 1966 Romanian President Nicolae Ceaucescu introduced pronatalist policies, outlawed abortion and contraception, and took measures to enforce the law. Mandatory pelvic examinations at places of employment were imposed on women of reproductive age. Informers for the security police were stationed in maternity hospitals. Doctors could be prosecuted for performing unauthorized abortions, and nurses were to make unannounced supervisory visits to new mothers to determine whether they were taking proper care of their infants.

The consequences of this policy and its enforcement are presented in Figures 1 through 3, through data from the Romanian birth and death registration system and the nationwide, ongoing maternal mortality audit system. (World Health Organization site visits from 1991 to 1992 have found that these systems are comparable to those of Western Europe in terms of the completeness of reporting and the reliability of data.[2] Romanian vital data systems and maternal mortality reports use ICD-9 CM definitions

Patricia Stephenson is with the Centrum for Folkhälsoforskning, Karlstad, Sweden. Marsden Wagner is retired from the World Health Organization, European Regional Office, Copenhagen, Denmark. Mihaela Badea is with the Ministry of Health of Romania, Bucharest, Romania. Florina Serbanescu is with the Institute for Mother and Child, Bucharest, Romania.

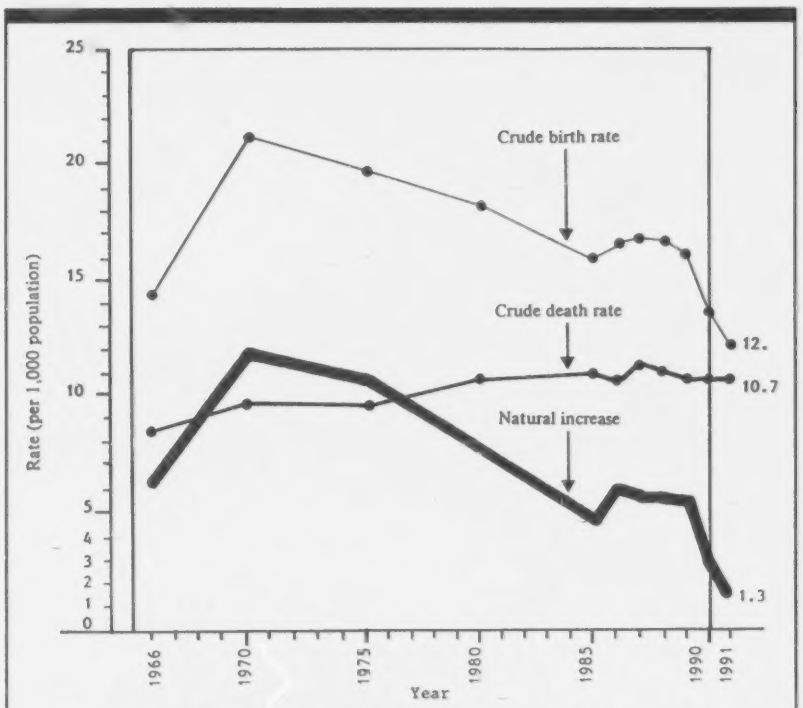
Requests for reprints should be sent to Patricia Stephenson, ScD, Centrum for Folkhälsoforskning, Landstinget i Värmland, 651 82 Karlstad, Sweden.

Editor's Note. See related editorial by Susser (p 1323) in this issue.

and diagnostic classifications.) After a brief rise, the crude birth rate fell and continued to fall (Figure 1). Thus the policy intended to increase the birth rate failed.

Before the 1966 law went into effect, the Romanian maternal mortality rate was similar to those of other Eastern European countries. Afterward, abortion-related maternal mortality increased to a level 10 times that of any other European country (Figures 2 and 3). For the decade 1980 to 1989, the average Romanian maternal mortality rate was 150 maternal deaths per 100 000 live births.[6] Many women obtained abortion illegally, and every year approximately 500 otherwise-healthy women of childbearing age died from postabortion hemorrhage, sepsis, abdominal trauma, and poisoning.

Precise figures on the Romanian prevalence of postabortion morbidity are not available. The country's most common causes of obstetrical death were postpartum hemorrhage and infection,[3] complications that could be associated with unsafe abortion. In Romania, local health professionals report that unsafe abortion—as performed by the woman herself or by untrained persons—may involve very dangerous techniques, such as scraping the uterus with a rubber tube or other instrument, uterine lavage with water or a caustic fluid, introduction of foreign bodies into the uterus, or external trauma. Such methods often result in damage to the uterine cervix, chronic infection, and severe anemia that, in turn, increase the risk of postpartum hemorrhage and infection, infertility, preterm birth, and low birthweight.[7,8]



Source. Adapted with permission from Ministry of Health of Romania.[3]

FIGURE 1—Crude birth rate, crude death rate, and rate of natural increase, Romania, 1966 to 1991.

The Social Impact of Restricted Abortion

Legions of Romanian children in institutions are another sad legacy of the years of the Ceaucescu pronatalist policy. Some women who did not have illegal abortions bore children after unwanted pregnancies and placed the children in institutions because they and their families simply could not find the means or motivation to care for them.[9] Although the number of children institutionalized before the overthrow of the Ceaucescu government is unknown, shortly after the revolution approximately 150 000 to 200 000 children were in institutional care.[10] Warehousing of children in institutions overwhelmed the health care system and reduced the standard of care.[9] Present programs have returned many institutionalized children to their families or have attempted to place them in adoptive homes.[9] However, several thousand severely retarded and handicapped children are likely to remain in institutions for the rest of their lives.[9]

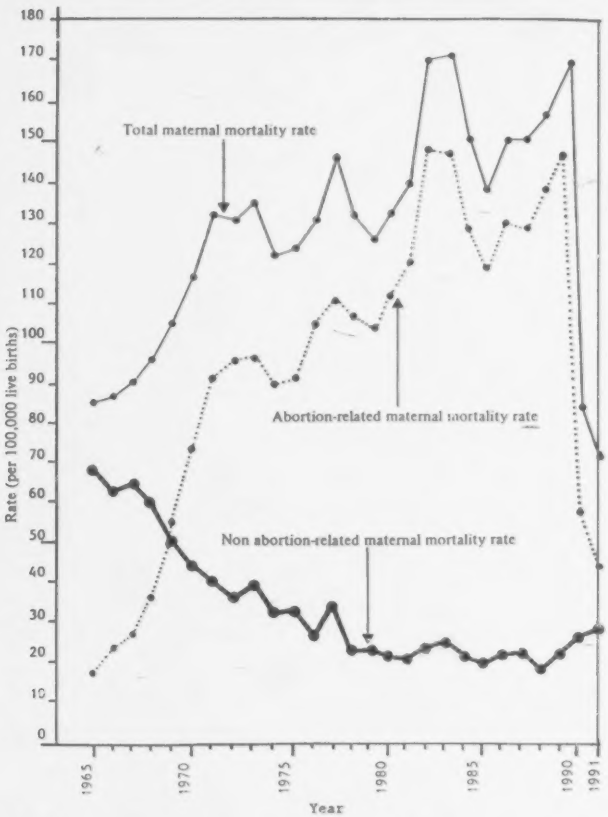
Conclusions

After the December 1989 Romanian revolution, one of the provisional government's first acts abrogated the 1966 law banning abortion and contraception. This was done as an emergency public health measure to try to decrease maternal mortality due to unsafe abortion. Since then, more and more induced abortions have been performed by qualified doctors in hospitals or clinics; the maternal mortality rate fell by 50% in the first year following the change in the law. It continues to fall as more and more women avail themselves of safe abortion. New admissions to children's institutions have decreased in spite of severe economic conditions.[9]

The Romanian experience demonstrates the futility and folly of attempts to control reproductive behavior through legislation. A law that forbids abortion does not stop women from aborting unwanted pregnancies. In Ireland today, where abortion is illegal, some 4000 women each year travel to Britain for the procedure. Nor can laws concerning reproductive behavior be effectively enforced. The extreme pronatalist policies in Romania did not succeed in sustaining the desired levels of reproduction and natural increase (Figure 1).

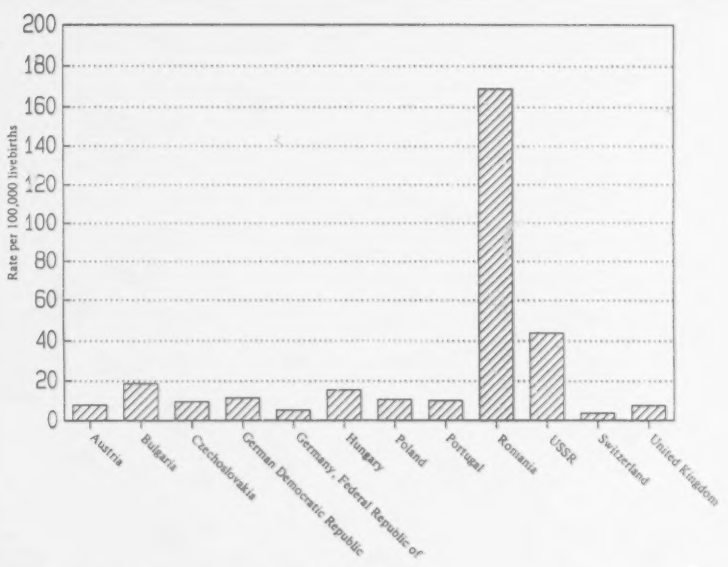
To reduce or eliminate abortion, the rational approach is to promote contraception. To paraphrase the position statement of the 1984 International Conference on Population, the solution to the abortion question is to prevent unwanted pregnancy.[7] The outlawing of abortion and contraception is not pronatalist but antinatalist in that the results are likely to be fewer healthy, fertile women who might have children, reduced fertility from widespread gynecological infection, the birth of more damaged babies, and, in sum, not more but fewer healthy citizens for the future.

In the 23 years of its enforcement, the antiabortion law in Romania resulted in over 10 000 deaths of women from unsafe abortion. The pronatalist policy as a whole resulted in the untoward institutionaliza-



Source. Adapted with permission from Ministry of Health of Romania.[3,4]

FIGURE 2—Abortion- and nonabortion-related maternal mortality rates and total maternal mortality rate, Romania, 1965 to 1991.



Source. Adapted from World Health Organization.[5]

FIGURE 3—1989 maternal mortality rate for 12 European and Asian countries.

tion of thousands of children and in an elevated rate of infant deaths. In effect, it also turned otherwise–law-abiding women and their partners into criminals; women learned to fear the government and its representatives. Doctors and other health care providers were victims as well because they were forced to carry out the policies of the state. Working today to improve and strengthen women's health services, Romanian doctors report that many women remain fearful of contact with the health care system.

The lessons of the Romanian experience hit home for those of us who remember how it was in our own countries prior to the legalization of abortion. Those who could afford to fly to other places could obtain legal, safe abortion, while those who could not made do with "backstreet" procedures or relinquished their unwanted children to orphanages. Perhaps we have come so far and accomplished so much that we have forgotten the public health consequences of restricting access to safe abortion.

In light of those consequences, it does not matter whether, as health professionals, we believe that life begins at conception or whether 8-week-old fetuses should have equal protection under the law. Policy on abortion is being formulated in the rarified atmosphere of legislatures and parliaments, meeting halls and court rooms. Let us hope that policymakers accept their responsibility for protecting the health and lives of women and children. □

References

1. United Kingdom Department of Health. *Report on Confidential Enquiries into Maternal Deaths in England and Wales, 1982–84. Report on Health and Social Subjects 34.* London, England: Her Majesty's Stationery Office; 1990.
2. World Health Organization, European Regional Office, Maternal and Child Health Unit. *Report of an MCH Mission to Romania.* Copenhagen, Denmark: World Health Organization; January 1991.
3. *Birth and Death Register.* Bucharest, Romania: Ministry of Health of Romania; 1992.
4. Maternal Mortality Audit System. Bucharest, Romania: Ministry of Health of Romania; 1992.
5. Health for All Database. Copenhagen, Denmark: World Health Organization, European Regional Office; 1991.
6. *Tabulated Birth and Death Statistics, 1991.* Bucharest, Romania: Ministry of Health of Romania; 1992.
7. World Health Organization, Maternal and Child Health Unit and Family Planning Division of Family Health. *Abortion: A Tabulation of Available Data on the Frequency and Mortality of Unsafe Abortion.*

Geneva, Switzerland: World Health Organization; 1990.
8. Royston E, Armstrong S. *Preventing Maternal Deaths*. Geneva, Switzerland: World Health Organization; 1989.
9. The Children's Health Care Collaborative Study Group. *The Causes of Institutionalization in Romanian Leagane and Sectii de Distrofici: Report of a Population-based Study with Recommendations*. Bucharest, Romania: Ministry of Health of Romania, Institute for Care of Mother and Child, and the United Nations Children's Fund; 1991.
10. *Report of a UNICEF Mission to Develop Emergency Assistance Programme for Institutionalized Children in Romania*. New York, NY: United Nations Children's Fund; 1990.

Public Health Policy Forum Submissions Invited

The "Public Health Policy Forum," a new feature of the Journal, is intended to provide a presentation of divergent views on important public health policy issues that is more extensive than the Journal's format usually allows. Three kinds of material will appear in the Public Health Policy Forum: articles, not to exceed 4500 words; commentaries, not to exceed 2500 words; and editorials. Customarily, commentaries and editorials will be solicited by the Journal's Associate Editor or Editor. Articles for the Public Health Policy Forum will be selected from articles submitted to the Journal and, on occasion, from those submitted directly to the Associate Editor. In general, peer review will still be sought.

Those wishing to submit articles directly to the Associate Editor may do so at this address:

Bruce C. Vladeck, PhD
The United Hospital Fund
55 Fifth Avenue
New York, NY 10003-4392

Five copies of the article should be submitted simultaneously to the Journal's Washington, DC, office.

The scholarly merit and the scientific accuracy of all submissions will be considered. Additional criteria are relevance to an important policy issue, timeliness, and clarity and coherence of the policy argument. Brevity also helps considerably.

Determinants of Depressive Symptoms in the Early Weeks after Miscarriage

Richard Neugebauer, PhD, MPH, Jennie Kline, PhD, Patricia O'Connor, PhD, Patrick Shrout, PhD, Jim Johnson, PhD, Andrew Skodol, MD, Judith Wicks, BA, and Mervyn Susser, MB, BCh, FRCP(E), DPH

ABSTRACT

Objectives. We tested whether and under what conditions miscarriage increases depressive symptoms in the early weeks following loss.

Methods. We interviewed 232 women within 4 weeks of miscarriage and 283 pregnant women and 318 community women who had not recently been pregnant. Depressive symptoms were measured with the Center for Epidemiologic Studies Depression (CES–D) Scale.

Results. Among women who had miscarried, the proportion who were highly symptomatic on the CES–D was 3.4 times that of pregnant women and 4.3 times that of community women. Among childless women, the proportion of women who had miscarried who were highly symptomatic was 5.7 times that of pregnant women and 11.0 times that of community women. Women who had miscarried were equally depressed regardless of length of gestation; among pregnant women, depressive symptoms declined with length of gestation. Among women who had miscarried, symptom levels did not vary with attitude toward the pregnancy; among pregnant women, depressive symptoms were elevated in those with unwanted pregnancies. Prior reproductive loss and advanced maternal age (35+ years) were not associated with symptom levels in any cohort.

Conclusions. Depressive symptoms are markedly increased in the early weeks following miscarriage. This effect is substantially modified by number of living children, length of gestation at loss, and attitude toward pregnancy. (*Am J Public Health.* 1992;82:1332–1339)

Introduction

Twelve to fifteen percent of clinically recognized pregnancies end in spontaneous abortion;[1] approximately half a million women miscarry annually in the United States alone.[2] For many women, miscarriage is an unanticipated and physically traumatic event signaling an abrupt disruption of reproductive plans. It may also prompt doubts about reproductive competence and thereby provoke a loss of self-esteem. Despite these considerations and repeated calls in the medical and lay literature for studies of the emotional effects of reproductive loss,[3–6] systematic observational studies of the psychological impact of miscarriage are rare. We have located only four published studies that focused exclusively on miscarriage,[7–10] and six that included miscarriage in an investigation of perinatal death.[11–16] In general, these studies are compromised by small sample sizes that limit both the generalizability of their results and the use of statistical adjustment for potentially confounding factors. More importantly, these investigations' lack of a comparison group unexposed to recent reproductive loss precludes determination of the overall size of the effect of miscarriage on psychiatric symptoms and the role of potentially moderating factors, such as maternal age and number of living children.

The frequency of miscarriage as an adverse reproductive outcome and the expectation that it will raise maternal risk for depressive symptoms makes assessment of its psychological impact an important public health question. We investigated whether and under what conditions miscarriage was associated with a substantial increase in depressive symptoms. In this paper we report results for the first few postabortion weeks. We hypothesized that miscarriage would raise the level of depressive symptoms in this early period after loss and that the magnitude of the effect would vary with characteristics of the woman, her pregnancy, her attitude toward the pregnancy, and her reproductive history. Specifically, we hypothesized that the effect of miscarriage would be greater in women with few or no living children, in those with a history of prior reproductive loss, in those with a positive attitude toward the pregnancy, and in those aged 35 years or older. Further, we hypothesized that the impact of miscarriage would increase with length of gestation at the time of loss. We tested these hypotheses by comparing levels of depressive symptoms in women who had miscarried with those in

Richard Neugebauer and Jennie Kline are with the New York State Psychiatric Institute and the Gertrude H. Sergievsky Center, Columbia University, New York, NY. Jennie Kline and Patrick Shrout are with the School of Public Health, Columbia University; Patrick Shrout is also with the Department of Psychology, New York University. Patricia O'Connor is with the Gertrude H. Sergievsky Center, as was Jim Johnson (who died after the study was completed); Patricia O'Connor is also with the Department of Psychology, Russell Sage College, Troy, NY. Andrew Skodol is with the Department of Psychiatry, Columbia University. Judith Wicks is with the New York State Psychiatric Institute. At the time of this study, Mervyn Susser was director of the Gertrude H. Sergievsky Center, Columbia University. He is now the Editor of this journal.

Requests for reprints should be sent to Richard Neugebauer, PhD, MPH, Epidemiology of Developmental Brain Disorders Department, New York State Psychiatric Institute, 722 W 168th St, New York, NY 10032.

This paper was submitted to the Journal May 29, 1991, and accepted with revisions May 22, 1992.

two other groups of women, both unexposed to recent reproductive loss.

Methods

Study Design

The study comprised three cohorts: women who had miscarried, currently pregnant women, and women in the community who had not been pregnant in the preceding year. The pregnant cohort afforded an estimate of the level of depressive symptoms to be expected in an uninterrupted pregnancy; the community cohort, the level of such symptoms in the absence of any recent reproductive event.

Miscarriage and Pregnant Cohort

In this study miscarriage is defined as the involuntary termination of a nonviable intrauterine pregnancy before 28 completed weeks of gestation. The miscarriage and pregnant cohorts were taken from the case subjects and control subjects, respectively, of a hospital-based case-control investigation of risk factors for miscarriage.[17] The case-control study attempted to interview all women attending a New York City hospital for a miscarriage (cases) from late 1984 to 1986. The control patients were drawn from women registered for prenatal care before 22 weeks' gestation and selected to be similar to case subjects in age and payment status (i.e., private or public).[18] The case-control study succeeded in interviewing 81% of identified case and control subjects. Among women who had miscarried and pregnant women, interviewed and uninterviewed subjects in the case-control study were distributed similarly on sociodemographic and reproductive characteristics.[17,19]

At the conclusion of the case-control interview, women who had miscarried and pregnant women who were 18 years old or older, English- or Spanish-speaking, and accessible by telephone were invited to enter the present investigation, which was separate from the case-control study. For pregnant women, we required further that the case-control interview occur before 28 completed weeks' gestation.

For the present study, we interviewed 232 women within 4 weeks of miscarriage; 85% were interviewed between days 7 and 15 after loss (Table 1). These women represented 70% of the eligible subjects then available from the case-control study. (Of the remainder, approximately half declined to participate altogether. The rest could not be scheduled for an interview within this narrow time frame but were evaluated in subsequent months. Together with later recruits from the case-control study, these subjects constitute material for a separate report.[20])

Seventy-two percent (283) of the eligible pregnant women were interviewed, most in their second trimester. Twenty percent refused to participate in the study; the remainder postponed the study interview beyond 28 weeks' gestation. Interviewing of women who had miscarried and pregnant women was concluded in 1987.

Community Cohort

Recruits for the community cohort were located by random dialing of telephone numbers. They were residentially stratified by the area codes and exchanges of interviewed women who had miscarried. A screening interview determined whether an eligible subject—a woman aged 18 to 44 years who had not been pregnant in the preceding 12 months—resided at the location. (Eligibility status was determined for 92% of working telephone numbers.) Community women were frequency-matched to the miscarriage cohort on language and season of interview, age (in 5-year intervals), and education (less than high school, high school graduate, college graduate, postgraduate training). Ethnicity was also recorded. Eighty-two percent (318) of known eligible community women were interviewed; the remainder declined to participate in the study. Community interviews were conducted in 1986 and 1987.

Comparability of Interviewed and Uninterviewed Study-Eligible Subjects

Within each cohort, interviewed and all uninterviewed eligible women were generally similar in age, ethnicity, education, and language and season of interview. Interviewed and uninterviewed women who had miscarried and pregnant women (on whom additional sociodemographic data were available from the case-control study) did not differ in marital status, parity, length of gestation or prior reproductive loss. However, as we will explain, public patients with miscarriages were given priority when in-hospital interviews in the case-control study were scheduled. As a result, among miscarrying subjects, women interviewed in the early weeks after loss tended to be disproportionately Black and Hispanic and less educated than uninterviewed women.

At the time of recruitment, women who had miscarried and pregnant women entering in the second half of the study, and all community women, were administered a version of an established single-item measure of depressed mood.[21] Within each cohort, the adjusted mood scores of interviewed and uninterviewed women were indistinguishable, suggesting that the affective state of the women successfully interviewed in our study was representative of study-eligible women generally.

Study Measures

Depressive symptoms were evaluated by means of the Center for Epidemiologic Studies Depression (CES–D) Scale,[22] a 20-item symptom checklist that asks about the presence and duration of symptoms in the preceding 7 days. This scale measures affective and somatic aspects of depressive symptomatology and low self-esteem. Scale items offer a fixed, alternative-choice format for subjects' responses regarding symptom frequency, scored from 0 ("rarely") to 3 ("most of the time"); item scores are summed to produce an overall index. The CES–D correlates well with other self-report measures of depressive symptoms, distinguishes clinically depressed patients from the general population,[22–24] and has been used previously with samples of pregnant women.[25] The CES–D scores of our subjects were associated with their scores on the single-item measure of mood administered several weeks earlier. Internal consistency reliabilities (Cronbach's alpha) of the CES–D in our three study cohorts overall and in language, ethnic, and educational subgroups within cohorts are excellent (.83 to .92).

We assessed attitudes toward the pregnancy in women who had miscarried and pregnant women with three items concerning the woman's wish to conceive, her emotional reaction to learning of her pregnancy, and her consideration of elective abortion. A subject's score could range from 0 to 12. A woman who had been trying to conceive, who reported being "very happy" when she learned she was pregnant, and who had given no thought to an elective abortion received a score of 12. A woman who did not want to become pregnant, who indicated being "very unhappy" about being pregnant, and who had actually scheduled an elective abortion before she miscarried scored 0. The internal consistency reliabilities of this scale in our two obstetric cohorts overall and in sociodemographic sub-

TABLE 1—Selected Sociodemographic and Reproductive History Characteristics of Women in Cohorts

	Cohort		
	Miscarriage (n = 232)	Pregnant (n = 283)	Community (n = 318)
Sociodemographic Characteristics			
Mean age, y (SD)	29.2 (6.3)	29.0 (5.7)	30.0 (6.4)
Ethnicity, %*			
White	25.4	36.4	36.6
Black	23.3	15.5	19.6
Hispanic	43.5	44.9	38.2
Other	7.8	3.2	5.7
Interviewed in Spanish, %	33.6	25.8	26.4
Education, %*			
< High school	30.6	24.7	20.3
High school graduate	26.7	20.9	24.5
Some college	23.7	25.8	29.7
College grad +	19.0	28.6	25.5
Marital status, %**			
Currently single	27.6	22.6	46.5
Married	57.8	60.8	35.2
Other	14.6	16.7	18.3
Hospital payment status**			
Private, %	28.4	39.2	...
Reproductive Characteristics			
Nulliparous, %**	35.3	34.6	50.5
No. living children, %**			
0	36.6	35.3	50.5
1	25.0	37.8	15.7
2	22.8	18.4	19.7
3+	15.5	8.5	14.1
No. prior reproductive losses, %[a]**			
0	63.8	71.0	82.5
1	21.6	20.8	12.9
2+	14.7	8.2	4.6
Length of gestation, %[b]**			
First trimester	62.5	1.4	...
Second trimester	37.5	83.7	...
Third trimester	...	14.8	...

Note. Differences among the three groups were evaluated by an overall chi-square test for categorical variables and by a one-way analysis of variance for continuous variables.
[a]Includes spontaneous abortions (89.6%), fetal deaths (3.7%), ectopic pregnancies (3.2%), and neonatal deaths (3.4%).
[b]For the miscarriage cohort, the gestation interval refers to the time from last menstrual period (LMP) to date of expulsion of the conceptus; for pregnant women, the interval is from LMP to the miscarriage study interview. First trimester = ≤ 13 completed weeks' gestation; second trimester = 14 through 26 completed weeks' gestation; third trimester = > 26 completed weeks' gestation.
*$P < .05$. **$P < .01$.

groups are acceptable (.68 to .78). More importantly, the scale correlated strongly with contraceptive use, pregnancy status, and reproductive plans at 6 months after loss, thereby affording evidence of its predictive validity in the miscarriage cohort and suggesting that subjects were not simply reporting positive, more socially acceptable attitudes toward their pregnancies.

Procedure

The CES–D, together with the remainder of the interview questions covering reproductive history, sociodemographic characteristics, and aspects of social functioning, was administered by telephone. The general comparability of sociodemographic, psychiatric symptom, and medical history data obtained from telephone and in-person interviews is well documented.[26-28]

The content of the interview precluded blinding interviewers to subjects' group membership. However, the possibility that interviewers influenced subjects' responses was minimized by the use of a fully structured questionnaire with fixed response options and monthly training sessions proscribing ad hoc interviewer probes. Compliance with this aspect of the study protocol was confirmed by periodic auditing of taped interviews by the first author.

Analytic Strategy

CES–D depressive symptom levels are reported in terms of symptom means for each cohort and in terms of proportions of women who were highly symptomatic. Women were considered highly symptomatic if they received CES–D scores of 30 or more, which meant that they reported having at least half of the 20 depressive symptoms most of the time for a week or longer. Analyses using these proportions assessed whether findings based on the entire range of symptom scores held true for extreme symptom levels. Approximately two thirds of the women scoring ≥30 would be expected to meet diagnostic criteria for major depressive disorder.[29]

For descriptive purposes, the observed CES–D means for each cohort overall and by selected characteristics are provided at the outset (Table 2). Comparison of means within cohorts is accomplished with one-way analyses of variance.

Formal analyses of differences in means and proportions across cohorts were conducted using ordinary least squares multiple regression and maximum likelihood logistic regression, respectively, to permit adjustment for potentially confounding variables. Candidate variables for adjustment in these multivariate analyses were drawn from an a priori list of potential confounding variables (e.g., marital status) and moderating variables (e.g., maternal age). (The matching variables for the community cohort were also introduced as covariates, both as main effects and as all possible two-way interaction effects. The latter terms did not materially influence the regression coefficients of interest here. Hence, for parsimony they are omitted from the final regression equations.) Screening for potential confounders in the data set (with reference to the

remaining characteristics, listed in Table 1) did not disclose other variables requiring analytic control. The regression equations comparing the women who miscarried and the community women contained the following terms: marital status, maternal age (linear), ethnicity, socioeconomic status (using a standardized index that assigns equal weight to maternal education and family income), number of living children, and prior reproductive loss. Equations comparing the women who miscarried and the pregnant women added payment status, length of gestation at miscarriage/interview (linear), and attitude toward the pregnancy.

Results from the least squares regression analyses are presented in the form of the difference in adjusted means between the miscarriage cohort and a comparison cohort; results from logistic regression analyses are presented as the ratio of the adjusted proportion of highly symptomatic women in the miscarriage cohort to the adjusted proportion of highly symptomatic women in a comparison cohort (Table 3).[30] (Corresponding odds ratios are provided in the notes to the relevant table.) The association of a particular characteristic, for example, number of living children, with depressive symptoms within each cohort is measured by the unstandardized regression coefficient *b* (see Figure 1 notes). The slope *b* indicates the change in CES-D score per unit change in the characteristic of interest. The possibility that the effect of miscarriage varies with the number of living children or, specifically in the miscarriage-pregnant comparison, with length of gestation or attitude toward pregnancy, was tested as a first-order interaction term in the regression analysis. Significant interaction terms were introduced as "centered" terms into all equations from the outset. For example, in calculating the adjusted difference in CES-D means between women who miscarried and pregnant women (Table 3), the estimated differences between cohorts were represented at the mean values for number of living children, length of gestation, and attitude toward the pregnancy.

Unless indicated otherwise, statistical significance is assessed at $P < .05$.

TABLE 2—Center for Epidemiologic Studies Depression Scale Means of Women in Cohorts, Overall and by Selected Characteristics

	Cohort					
	Miscarriage (n = 232)		Pregnant (n = 283)		Community (n = 318)	
	Mean	(SE)	Mean	(SE)	Mean	(SE)
Overall	24.0	(0.8)	15.2	(0.7)	14.6	(0.8)
Sociodemographic Characteristics						
Age, y						
18–24	26.7	(1.5)*	18.0	(1.4)**	14.3	(1.3)
25–34	23.2	(1.2)	15.0	(0.8)	15.0	(0.9)
35+	22.3	(1.6)	12.8	(1.6)	13.9	(1.3)
Education						
< High school	25.0	(1.5)*	18.1	(1.4)****	20.7	(1.6)****
High school graduate	25.5	(1.7)	17.9	(1.7)	18.1	(1.4)
Some college	23.8	(1.7)	15.7	(1.2)	11.7	(0.9)
College grad +	20.3	(1.6)	10.4	(1.0)	11.0	(1.1)
Marital status						
Married	22.3	(1.1)*	12.4	(0.7)****	12.4	(1.0)***
Currently single	26.4	(1.6)	20.4	(1.5)	14.6	(0.9)
Other	25.8	(2.0)	18.6	(1.9)	18.6	(1.5)
Hospital payment status						
Private	19.9	(1.4)	10.0	(0.8)	...	
Public	25.6	(1.0)	18.6	(0.9)	...	
Reproductive History Characteristics						
No. living children						
0	26.0	(1.3)**	12.7	(1.0)***	13.6	(0.8)***
1	25.0	(1.6)	14.9	(1.0)	14.6	(1.8)
2	20.7	(1.6)	20.0	(1.8)	14.4	(1.3)
3+	22.1	(2.3)	16.7	(2.6)	20.2	(1.9)
No. prior reproductive losses						
0	23.9	(1.0)	15.1	(0.8)	14.3	(0.7)
1	25.8	(1.8)	15.1	(1.4)	18.6	(2.1)
2+	21.6	(2.3)	16.1	(2.3)	13.6	(3.5)
Length of gestation[a]						
First trimester	23.4	(1.0)	20.0	(5.1)*	...	
Second trimester	24.8	(1.3)	15.6	(0.7)	...	
Third trimester	...		12.8	(1.6)	...	
Attitude toward pregnancy[b]						
Wanted	24.1	(1.8)	11.1	(0.9)****	...	
Mixed	24.1	(1.0)	15.1	(0.9)	...	
Unwanted	23.1	(2.2)	25.6	(1.9)	...	

Note. *P* values are derived from ANOVA tests for within-cohort differences in means. For age, education, number of living children, number of prior reproductive losses, length of gestation, and attitude toward pregnancy, *P* values pertain to two-tailed ANOVA tests for linearity, using 1 degree of freedom.

[a]First trimester = ≤ 13 completed weeks' gestation; second trimester = 14 through 26 completed weeks' gestation; third trimester = > 26 completed weeks' gestation.

[b]In both the pregnant and miscarriage cohorts, roughly 15% of the women indicated unequivocally that they did not want the pregnancy. Approximately 55% indicated exclusively positive feelings of wanting the pregnancy, and the remainder expressed some reservations about their pregnancies.

*$P < .10$; **$P < .05$; ***$P < .01$; ****$P < .001$.

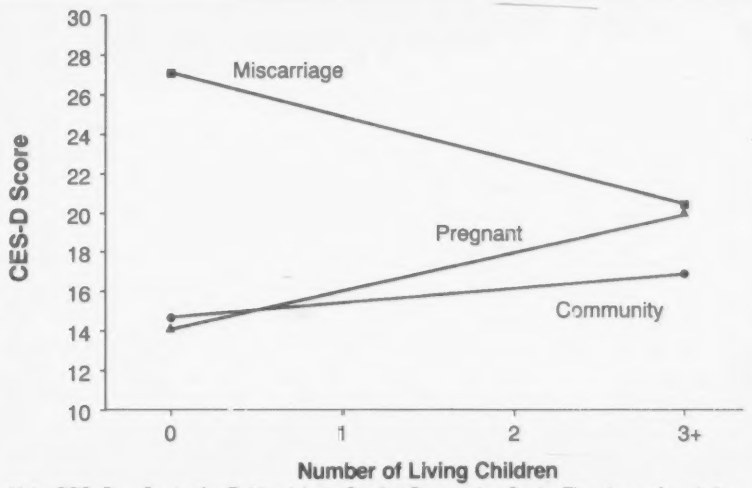
Sample Characteristics

Community women were selected to be sociodemographically similar to all interviewed women who had miscarried, including those assessed only in later months of the study. However, the case-control study gave priority to conducting in-hospital interviews with miscarrying women who were public patients; private patients were judged easier to reach after discharge. As a consequence, public patients are overrepresented among women interviewed in the present study in the early postabortion weeks. Therefore, the miscarriage cohort in this report differs from the other two groups in those characteristics associated with payment status, for example, ethnicity and educational level (Table 1). Also, as expected, the two obstetric cohorts contain a lower

TABLE 3—Center for Epidemiologic Studies Depression Scale (CES–D) Means and Proportions of Highly Symptomatic Women

	Miscarriage Cohort (n = 232)	Pregnant Cohort (n = 283)	Community Cohort (n = 318)
Means			
Observed CES–D mean (± SE)	24.0 (± 0.8)	15.2 (± 0.7)	14.6 (± 0.8)
Adjusted difference in means[a] (95% CI) between the miscarriage cohort and this cohort		7.5 (4.5, 10.5)	9.1 (6.9, 11.2)
Proportions			
Observed % scoring ≥ 30 (± SE)	36.2 (± 3.2)	12.4 (± 2.0)	10.1 (± 1.7)
Adjusted ratio of proportions[a] (95% CI) of highly symptomatic women in the miscarriage cohort to proportions of highly symptomatic women in this cohort		3.4 (2.0, 5.0)[b]	4.3 (3.0, 5.8)[c]

[a]Means are adjusted by ordinary least squares multiple regression analysis, proportions by maximum likelihood logistic regression analysis.
[b]The odds ratio for this comparison is 4.9 (95% CI = 2.4, 10.0).
[c]The odds ratio for this comparison is 6.2 (95% CI = 3.8, 10.5).



Note. CES–D = Center for Epidemiologic Studies Depression Scale. The slope of each line (*b*) represents the unstandardized regression coefficient for the term for number of children, from within-cohort least squares multiple regression. In the pregnant cohort, $b = 1.3$ per child (95% CI = −0.1, 2.8), $P < .11$; in the miscarriage cohort, $b = -2.1$ per child (95% CI = −3.8, −0.3), $P < .02$; in the community cohort, $b = 0.73$ per child (95% CI = −0.72, 2.2), $P < .32$. The interaction term in the pregnant/miscarriage analysis (cohort by number of children) is significant at $P < .002$. The interaction term in the community/miscarriage analysis is significant at $P < .0004$.

FIGURE 1—Number of children and adjusted CES–D scores in miscarriage, pregnant, and community cohorts.

proportion of nulliparous women and a higher proportion of women with prior reproductive loss than does the community cohort. The regression analyses adjust satisfactorily for the possible effect of these sociodemographic and reproductive history differences between cohorts on depressive symptoms.

Results

Overall CES–D Means and Proportions of Women Highly Symptomatic

Depressive symptom levels were higher in women who had miscarried than in women in the pregnant and community cohorts. The observed symptom mean of the miscarriage cohort was 24.0 (Table 2); 36.2% of the subjects were highly symptomatic (Table 3). In adjusted analyses, the mean of the miscarriage cohort exceeded those of the pregnant and community cohorts by 7.5 and 9.1 points, respectively (Table 3). Similarly, in adjusted analyses, the proportion of subjects highly symptomatic in the miscarriage cohort was 3.4 and 4.3 times the proportion in the pregnant and community cohorts, respectively (Table 3).

Number of Living Children

The presence of children strongly influenced the effect of miscarriage on depressive symptoms. In the pregnant and community cohorts, the observed symptom means exhibit a significant increase with an increase in the number of children. By contrast, in the miscarriage cohort, depressive symptoms declined with increasing numbers of children (Table 2). These differences are sustained in multivariate analyses. The slopes of the regression lines representing the association of numbers of children with depressive symptoms differ markedly between the pregnant and miscarriage cohorts (Figure 1). Therefore, despite the overall difference in symptom levels between pregnant women and women who had miscarried, average symptom levels did not differ between miscarrying and pregnant women with several children. The same pattern holds in the comparison of the miscarriage and community cohorts (Figure 1).

To delineate further an especially high-risk group, we dichotomized our subjects into those with and without children. Forty percent of the childless women who miscarried were highly symptomatic. In adjusted analyses, this proportion is 5.7 times that of childless pregnant women (95% CI = 2.5, 8.9) and 11 times that of childless community women (95% CI = 5.6, 16.7). Among women with children, the proportion highly symptomatic was 34.0 for women who had miscarried;

this proportion is 2.9 times that of pregnant women with children (95% CI = 1.5, 5.1) and 2.3 times that of community women (95% CI = 1.5, 3.3) in adjusted analyses.

Length of Gestation

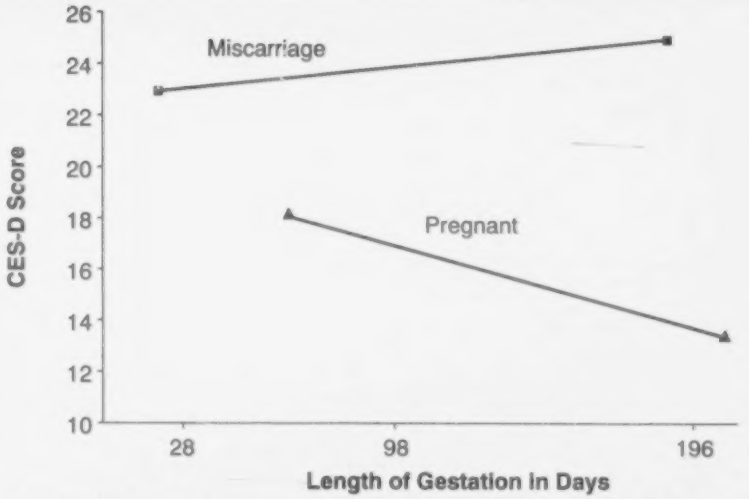
The association of length of gestation with depressive symptoms differs between the miscarriage and pregnant cohorts. In the pregnant cohort, observed symptom means declined with gestational advance (Table 2). In the miscarriage cohort, depressive symptoms were not associated with length of gestation; that is, women with early and late miscarriages had equally high levels of depressive symptoms. Hence, given the decline in symptom levels expected at later stages of an uninterrupted pregnancy, miscarriage is seen to exert a greater effect on women at later points in gestation (Figure 2). (This pattern of results persisted even after we restricted the analysis to gestational periods characterized by greater overlap between the two cohorts.)

Attitude toward the Pregnancy

The association of depressive symptoms with attitude toward pregnancy also varied between the miscarriage and pregnant cohorts. In the pregnant cohort, depressive symptom levels were higher in women with unwanted pregnancies; in the miscarriage cohort, depressive symptom levels were unrelated to the woman's attitude toward the pregnancy (Table 2, Figure 3). Hence, when pregnant women are used as the standard, miscarriage has a greater impact on women with wanted pregnancies than on women with unwanted pregnancies. However, these effects of attitude do not account for the pattern of results pertaining to number of children, described earlier, since that pattern persists when the analysis is restricted to women with wanted pregnancies.

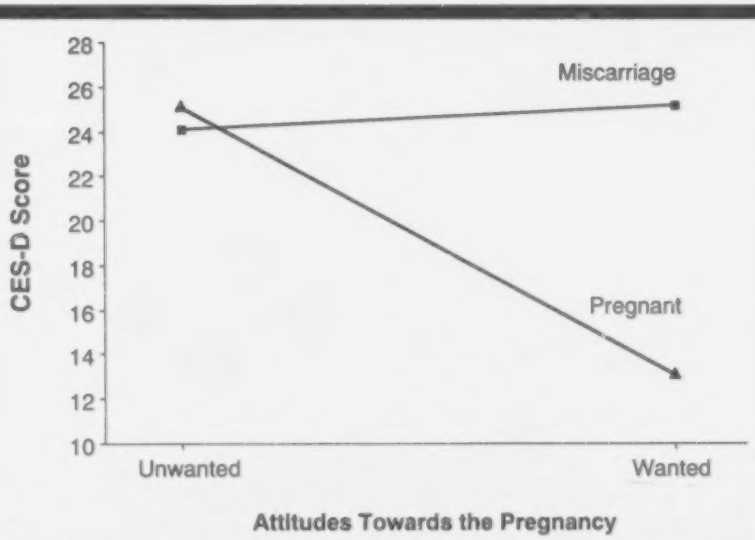
Prior Reproductive Loss, Maternal Age, Marital and Socioeconomic Status, and Religion

Depressive symptom levels were not raised in women with prior reproductive loss or in older women (aged 35 and older) within any study cohort in either uncontrolled (Table 2) or controlled analyses. Hence, miscarriage does not appear to exert a disproportionately greater effect on women with prior reproductive loss or on older women. These results did not change in analyses confined to nulliparous women. The effect of miscarriage on depressive symptoms also did not vary by marital status.



Note. CES-D = Center for Epidemiologic Studies Depression Scale. The slope of each line (*b*) represents the unstandardized regression coefficient for the term for length of gestation in days, from within-cohort least squares multiple regression equations. In the pregnant cohort, $b = -0.04$ per day (95% CI = −0.09, 0.008), $P < .11$; in the miscarriage cohort, $b = .02$ per day (95% CI = −0.03, 0.06), $P < .46$. The interaction term (cohort by length of gestation) is significant at $P < .08$.

FIGURE 2—Length of gestation and adjusted CES-D scores in miscarriage and pregnant cohorts.



Note. CES-D = Center for Epidemiologic Studies Depression Scale. The slope of each line (*b*) represents the unstandardized regression coefficient for the term for pregnancy attitude, from within-cohort least squares multiple regression equations. In the pregnant cohort, $b = -1.05$ (95% CI = −1.5, −0.6), $P < .00001$; in the miscarriage cohort, $b = 0.07$ (95% CI = −0.6, 0.7), $P < .82$. The interaction term (cohort by attitude) is significant at $P < .009$.

FIGURE 3—Attitude toward the pregnancy and adjusted CES-D scores in the miscarriage and pregnant cohorts.

Depressive Symptoms vs Somatic Complaints

The depressive symptoms reported by subjects in our study might have consisted largely of somatic complaints and affective disturbances arising from the physical sequelae of the miscarriage (or of the pregnancy). To evaluate this possibility, we repeated all of the preceding analyses after

first removing the somatic items (e.g., appetite and sleep disturbance) from the CES–D and then entering the summed score of the deleted items as a covariate in the regression equations. The preceding patterns persisted. Furthermore, although one third of the women who had miscarried had been hospitalized overnight and another third had been hospitalized for two or more nights, hospital stay was unassociated with symptom levels. Women who had had a dilation and curettage (approximately 85%) and those who had not also did not differ on level of depressive symptoms.

Discussion

Depressive symptoms were increased substantially in the early weeks after loss in this study sample. Symptom means in the miscarriage cohort approximated levels reported for heterogeneous samples of psychiatric inpatients.[22] The proportion of women who were highly symptomatic was 3.4 and 4.3 times greater in the miscarriage cohort than in the pregnant and community cohorts, respectively.

Childless women who had miscarried proved especially vulnerable to depressive symptoms. For example, the proportion highly symptomatic was 11 times higher among childless women who had miscarried than among childless community women. By contrast, the presence of several children is entirely protective. Several prior investigations of miscarriage, albeit not all,[9] have also reported lower symptom levels with increasing numbers of children.[11,15,31] However, since these studies lacked comparison groups they did not disclose the more striking finding that women with several children who miscarry do not exhibit significantly higher symptom levels than women with children who have not experienced a recent reproductive loss.

The apparent absence of a depressive effect of miscarriage on women with several children contradicts a large body of research linking life events involving loss to increases in depressive symptoms.[32] Recent stress research has focused on factors that buffer individuals from the pathogenic effects of negative life events, with much attention accorded to social supports.[33] For a woman who miscarries, the presence of living children may afford psychological support indirectly, by representing evidence of reproductive success in the past. This finding merits further investigation.

We found women with early and late miscarriages to be equally depressed, as has been noted in some other studies.[7,12,13] However, compared with pregnant women of similar gestational lengths, women with late loss apparently experienced a greater rise in depressive symptoms than women with earlier loss. This finding of a greater effect of miscarriage with length of gestation is consistent with the notion that maternal attachment to the unborn child progresses as the pregnancy advances and that the impact of loss corresponds to the strength of the preceding attachment.[31,34]

Again, when pregnant women are used as a reference group, loss of a wanted pregnancy is associated with a marked increase in symptom levels, whereas loss of an unwanted pregnancy is not. Women with unwanted pregnancies already have elevated symptoms and therefore apparently experience no further increase at the time of miscarriage. However, women losing wanted and unwanted pregnancies are equally depressed, as reported elsewhere.[8,15,31] Both situations warrant clinical and public health concern, as do pregnant women carrying unwanted conceptions.

Women with a history of prior reproductive loss, consisting predominantly of miscarriages, were not affected more strongly by their miscarriages than were women without prior loss. Previous studies of miscarriage[11,13,14,31] found no association between prior loss and current symptoms in their study cohorts. With one exception,[35] findings from previous investigations of perinatal death concur.[16,36–39] Our study findings also obtained in analyses restricted to nulliparous subjects.

With the current trend in industrialized societies of women postponing childbearing until later years, the effect of miscarriage on older women deserves special mention. Our finding that older women in the miscarriage cohort were not more depressed than younger women agrees with earlier results from research on women who had miscarried[10,11,13,14,31] and on women experiencing perinatal death.[16,36–37,39] Restriction of our analyses to nulliparous women did not alter these findings.

This study demonstrates for the first time that miscarriage powerfully influences maternal psychiatric status in the first few weeks after loss. Physicians in general practice, as well as gynecologists, should be aware both that nulliparous women may be in special need of attention following a miscarriage and that women experiencing even very early loss or the loss of an ostensibly unwanted pregnancy may exhibit markedly high levels of depressive symptoms. The possible persistence of these effects among certain subgroups of women awaits future investigation. □

References

1. Kline J, Stein Z, Susser M. *Conception to Birth: Epidemiology of Prenatal Development.* New York, NY: Oxford University Press; 1989:43–68.
2. *Vital Statistics of the United States 1986. Vol I. Natality.* Hyattsville, Md: National Center for Health Statistics; 1988.
3. Lewis E. The management of stillbirth: coping with an unreality. *Lancet.* 1976;19: 619–620.
4. Elliott BA, Hein HA. Neonatal death: reflections for physicians. *Pediatrics.* 1978; 62:96–100.
5. Zeanah CH. Adaptation following perinatal loss: a critical review. *J Am Acad Child Adolesc Psychiatry.* 1989;28:467–480.
6. Borg S, Lasker J. *When Pregnancy Fails: Families Coping with Miscarriage, Stillbirth and Infant Death.* Boston, Mass: Beacon Press; 1981.
7. Leppert PC, Pahlka BS. Grieving characteristics after spontaneous abortion: a management approach. *Obstet Gynecol.* 1984; 64:119–122.
8. Seibel M, Graves WL. The psychological implications of spontaneous abortions. *J Reprod Med.* 1980;25:161–165.
9. Friedman T, Gath D. The psychiatric consequences of spontaneous abortion. *Br J Psychiatry.* 1989;155:810–813.
10. Garel M, Blondel B, Lelong N, Papin C, Bonenfant S, Kaminski M. Reactions depressives après une fausse-couche. *Contracept Fertil Sex.* 1992;20:75–81.
11. Toedter LJ, Lasker JN, Alhadeff JM. The Perinatal Grief Scale: development and initial validation. *Am J Orthopsychiatry.* 1988;58:435–449.
12. Peppers LG, Knapp RJ. Maternal reactions to involuntary fetal/infant death. *Psychiatry.* 1980;43:155–159.
13. Nicol MT, Tompkins JR, Campbell NA, Syme GJ. Maternal grieving response after perinatal death. *Med J Aust.* 1986;144:287–289.
14. LaRoche C, Lalinec-Michaud M, Engelsmann F, et al. Grief reactions to perinatal death—a follow-up study. *Can J Psychiatry.* 1984;29:14–19.
15. Graham MA, Thompson SC, Estrada M, Yonekura ML. Factors affecting psychological adjustment to a fetal death. *Am J Obstet Gynecol.* 1987;157:254–257.
16. Rowe J, Clyman R, Green C, Mikkelsen C, Haight J, Ataide L. Follow-up of families who experience a perinatal death. *Pediatrics.* 1978;62:166–170.
17. Kline J, Stein Z, Susser M, Warburton D. Induced abortion and the chromosomal characteristics of subsequent miscarriages (spontaneous abortions). *Am J Epidemiol.* 1986;123:1066–1079.
18. Kline J, Hutzler M, Levin B, Stein Z, Susser M, Warburton D. Marijuana and spontaneous abortion of known karyotype. *Paediatr Perinat Epidemiol.* 1991;5:320–332.
19. Neugebauer R. The psychiatric effects of miscarriage: research design and prelimi-

nary findings. In: Cooper B, ed. *Psychiatric Epidemiology*. London: Croom Helm; 1987.

20. Neugebauer R, Kline J, O'Connor P, et al. Depressive symptoms in women in the six months following miscarriage. *Am J Obstet Gynecol.* 1992;166:104–109.
21. Folstein MF, Luria R. Reliability, validity and clinical application of the Visual Analogue Mood Scale. *Psychol Med.* 1973;3: 479–486.
22. Radloff LS. The CES–D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1:385–401.
23. Myers JM, Weissman MM. Use of a self-report symptom scale to detect depression in a community sample. *Am J Psychiatry.* 1980;137:1081–1084.
24. Weissman MM, Sholomskas D, Pottenger M, Prusoff BA, Locke BZ. Assessing depressive symptoms in five psychiatric populations: a validation study. *Am J Epidemiol.* 1977;106:203–213.
25. Zuckerman B, Amaro H, Bauchner H, Cabral H. Depressive symptoms during pregnancy: relationship to poor health behaviors. *Am J Obstet Gynecol.* 1989;160: 1107–1111.
26. Aneshensel CS, Frerichs RR, Clark VA, Yokopenic PA. Measuring depression in the community: a comparison of telephone and personal interviews. *Public Opinion Q.* 1982;46:110–121.
27. Leimbach R. Alternatives to the face-to-face interview for collecting gerontological needs assessment data. *Gerontologist.* 1982;22:78–82.
28. Hochstim JR. A critical comparison of three strategies of collecting data from households. *Am Stat Assoc J.* 1967; September:976–989.
29. Boyd JH, Weissman MN, Thompson D, Myers JK. Screening for depression in a community sample. *Arch Gen Psychiatry.* 1982;39:1195–1200.
30. Kelsey JL, Thompson WD, Evans AS. *Methods in Observational Epidemiology.* New York, NY: Oxford University Press; 1986.
31. Kirkley-Best E. Grief in response to prenatal loss: an argument for the earliest maternal attachment. Gainesville, Fla: University of Florida; 1981. Thesis.
32. Osterweiss M, Solomon F, Green M, eds. *Bereavement: Reactions, Consequences and Care.* Washington, DC: National Academy Press; 1984.
33. Cohen S, Syme LE, eds. *Social Support and Health.* New York, NY: Academic Press; 1985.
34. Condon JT. The parental-foetal relationship—a comparison of male and female expectant parents. *J Psychosom Obstet Gynecol.* 1985;4:271–284.
35. Kennell JH, Slyter H, Klaus MH. The mourning response of parents to the death of a newborn infant. *N Engl J Med.* 1970; 283:344–349.
36. Cullberg J. Mental reactions of women to perinatal death. In: *Proceedings of the Third International Congress on Psychosomatic Medicine in Obstetrics and Gynaecology.* Basel, Switzerland: Karger; 1972.
37. Benfield DG, Leib SA, Vollman JH. Grief response of parents to neonatal death and parent participation in decision care. *Pediatrics.* 1978;62:171–177.
38. Forrest GC, Standish E, Baum JD. Support after perinatal death: a study of support and counselling after perinatal bereavement. *Br Med J [Clin Res].* 1982;285: 1475–1479.
39. Tudehope DI, Iredall J, Rodgers D, Gunn A. Neonatal death: grieving families. *Med J Aust.* 1986;144:290–292.

Trends in Obstetric Operative Procedures, 1980 to 1987

S. Christine Zahniser, RN, MPH, Juliette S. Kendrick, MD, Adele L. Franks, MD, and Audrey F. Saftlas, PhD

ABSTRACT

Objectives. Increasing rates of cesarean deliveries have received widespread attention in recent years, as concern in the United States about unnecessary surgical procedures has increased. However, little information has been published on the national trends of other operative obstetric procedures occurring during deliveries.

Methods. We analyzed data from the National Hospital Discharge Survey to examine trends in the use of forceps, vacuum extraction, and cesarean section from 1980 through 1987.

Results. The rate of cesarean sections increased by 48%, while the rate of forceps procedures declined by 43%. Although the risk of cesarean section was significantly increased for older women, the risk of forceps and vacuum extraction procedures did not vary by age. Women with private insurance were significantly more likely to receive a cesarean section (rate ratio [RR] = 1.2), forceps procedure (RR = 1.7), and vacuum extraction procedure (RR = 1.8) than were women without private insurance.

Conclusions. As pressure mounts to decrease the national cesarean section rate from 24% to 15% by the year 2000, attention should also be given to surveillance of other operative delivery procedures. (*Am J Public Health.* 1992;82:1340–1344)

Introduction

Increasing rates of cesarean deliveries have received widespread attention in recent years, as concern in the United States about unnecessary surgical procedures has increased. In 1987, cesarean deliveries became the most frequently performed surgical procedure among women of childbearing years.[1] Some researchers have hypothesized that this increase in cesarean sections is the result of a decrease in the use of forceps.[2] However, limited information has been published on the national trends of other operative obstetric procedures occurring during deliveries.[3] Here, we examine US trends in the use of forceps, vacuum extraction, and cesarean section over an 8-year period. In addition, we present the demographic characteristics of women most likely to receive these interventions.

Methods

We analyzed data from the National Hospital Discharge Survey, conducted by the National Center for Health Statistics. Each year, a representative sample of discharges is selected from approximately 400 short-term–stay, nonfederal hospitals throughout the United States.[4] About 200 000 abstracts of medical records are sampled annually. Data are transcribed from the face sheet of the medical record to the National Hospital Discharge Survey abstract form. Abstracted items include demographic characteristics of the patient, administrative information, expected sources of payment, up to seven discharge diagnoses, and up to four procedures. Coding is performed according to the *International Classification of Diseases,* 9th revision, Clinical Modification (ICD-9-CM).[5]

We restricted our analysis to women hospitalized for a delivery (ICD-9-CM codes V27.0 through V27.9) from 1980 through 1987. The obstetric procedures we analyzed included the use of forceps (ICD-9-CM codes 72.0 through 72.49, 72.51, 72.53, 72.6, 73.3), vacuum extraction (72.7), and cesarean sections (74.0 through 74.29, 74.4, 74.99). For each delivery, we selected only one operative procedure code, even if more were present. For example, in the many cases in which both low forceps use (72.0) and forceps rotation of the fetal head (72.4) were recorded, we analyzed this combination as one forceps procedure.

We classified the mother's race as "White" or "Black and other minorities." Insurance status was established on the basis of information collected from patients at admission to the hospital. Women whose expected source of payment included any type of private or commercial insurance were reported as having "private insurance." Women who reported self-payment or government payment, including Medicaid/Medicare, were considered to have "no private in-

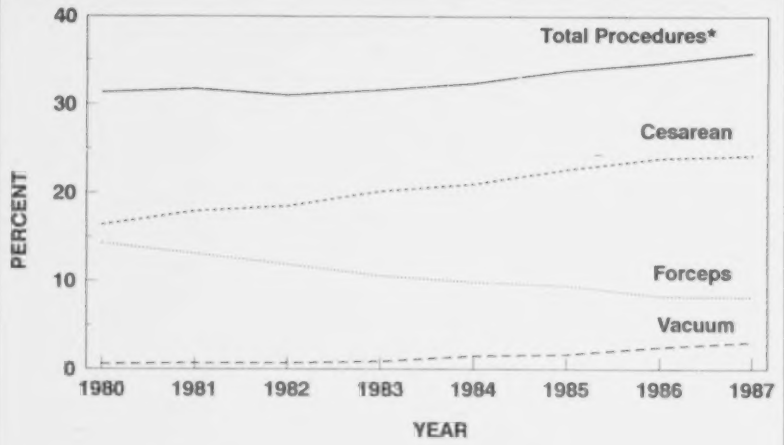
S. Christine Zahniser, Juliette S. Kendrick, and Audrey F. Saftlas are with the Division of Reproductive Health, and Adele L. Franks is with the Office of Surveillance and Analysis, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control, Atlanta, Ga.

Requests for reprints should be sent to S. Christine Zahniser, RN, MPH, Division of Reproductive Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control, Atlanta, GA 30333.

This paper was submitted to the Journal October 28, 1991, and accepted with revisions May 19, 1992.

surance." Women in the "other" payment group, which consisted of those in the "workman's compensation," "no charge," or "other" payment categories, represented less than 5% of the sample and were excluded from the analysis of insurance data. Rate ratio analyses of the 1987 deliveries were conducted to identify women at increased risk of having the various procedures.

We used the computer program SESUDAAN (Research Triangle Institute, Research Triangle Park, NC) to calculate rates and their corresponding standard errors,[6] and we took into account the complex multistage sampling design of the National Hospital Discharge Survey. We used the Taylor series expansion with the logit transformation to calculate 95% confidence intervals on rate ratios.[7] Weighted linear regression was used to test for linear trend in the rates of various procedures over the 8-year study period.[8] Rates and rate ratios are provided only for procedures with 5000 or more discharges; estimates of less than 5000 in the National Hospital Discharge Survey are considered statistically unreliable[3] since they are based on fewer than 30 records. We calculated crude and adjusted rates of cesarean section, forceps, and vacuum extraction procedures for 1987. Because the age- and race-adjusted rates were similar to the crude rates, we present only the crude rates here.



* Total procedures include cesarean section, forceps, and vacuum extraction

FIGURE 1—Rate of US obstetric operative procedures, 1980 to 1987, National Hospital Discharge Survey, National Center for Health Statistics, Hyattsville, Md.

Results

General Trends

The number of operative delivery procedures, including cesarean sections, forceps, and vacuum extraction procedures, increased from 1.2 million in 1980 to 1.4 million in 1987, affecting 35.8% of all deliveries in 1987, compared with 31.3% in 1980 (Figure 1). Between 1980 and 1987, cesarean sections increased by 48% (from 617 000 to 949 000), whereas forceps procedures declined by 43% (from 539 000 to 320 000). In comparison, vacuum extraction procedures increased from 0.6% to 3.3% of all deliveries (from 22 000 to 122 000); most of this increase occurred after 1983.

TABLE 1—Incidence Rates, Rate Ratios, and Confidence Intervals (CI) for US Obstetric Operative Procedures, by Selected Demographic Characteristics, 1987

	Cesarean Sections			Forceps Procedures			Vacuum Procedures		
	Rate[a]	Rate Ratio	95% CI	Rate[a]	Rate Ratio	95% CI	Rate[a]	Rate Ratio	95% CI
Age, y									
15–19	18.6	0.8	0.7, 0.8	8.2	1.0	0.7, 1.4	3.6	1.2	0.8, 1.9
20–34	24.5	1.0	Referent	8.2	1.0	Referent	3.0	1.0	Referent
35–44	31.6	1.3	1.2, 1.4	7.1	0.9	0.6, 1.2	3.7	1.2	0.7, 2.1
Race[b]									
White	25.1	1.1	1.0, 1.2	9.1	1.6	1.1, 2.3	3.2	1.4	0.8, 2.6
Black and other minorities	23.4	1.0	Referent	5.8	1.0	Referent	2.2	1.0	Referent
Insurance status									
Private insurance	25.6	1.2	1.1, 1.2	9.9	1.7	1.3, 2.3	3.8	1.8	1.2, 2.8
No private insurance	22.1	1.0	Referent	5.7	1.0	Referent	2.2	1.0	Referent
Geographic region									
Northeast	26.3	1.2	1.0, 1.3	6.0	1.2	0.8, 1.8	2.6	0.7	0.4, 1.4
Midwest	23.2	1.0	0.9, 1.2	7.6	1.5	1.0, 2.1	2.4	0.6	0.3, 1.3
South	25.4	1.1	1.0, 1.3	12.1	2.3	1.4, 3.8	3.5	0.9	0.5, 1.9
West	22.5	1.0	Referent	5.2	1.0	Referent	3.7	1.0	Referent

Note. Incidence rates, rate ratios, and confidence intervals were calculated using data from the National Hospital Discharge Survey, 1987, National Center for Health Statistics, Hyattsville, Md.

[a] Per 100 deliveries.

[b] Women with unknown race were excluded (13% of women with cesarean section, 14% of women with forceps procedures, and 19% of women with a vacuum extraction procedure).

TABLE 2—Rates of US Obstetric Operative Procedures, by Race and Insurance Status, 1987

	Private Insurance						No Private Insurance					
	Cesarean		Forceps		Vacuum		Cesarean		Forceps		Vacuum	
Race[a]	Rate	(95% CI)	Rate	(95% CI)	Rate	(95% CI)	Rate	(95% CI)	Rate	(95% CI)	Rate	(95% CI)
White	26.3	(24.9, 27.7)	10.3	(7.5, 13.2)	3.7	(2.7, 4.7)	22.4	(20.2, 24.5)	6.3	(5.0, 7.7)	2.3	(1.3, 3.2)
Black and other minorities	25.6	(23.0, 28.3)	7.1	(4.6, 9.7)	3.3	(1.1, 5.5)	21.9	(19.9, 23.8)	5.2	(3.8, 6.6)	1.4	(0.7, 2.1)

Note. Rates and confidence intervals were calculated using data from the National Hospital Discharge Survey, 1987, National Center for Health Statistics, Hyattsville, Md.
[a]Women with unknown race were excluded (13.5% of those with private insurance and 14.1% of those without private insurance).

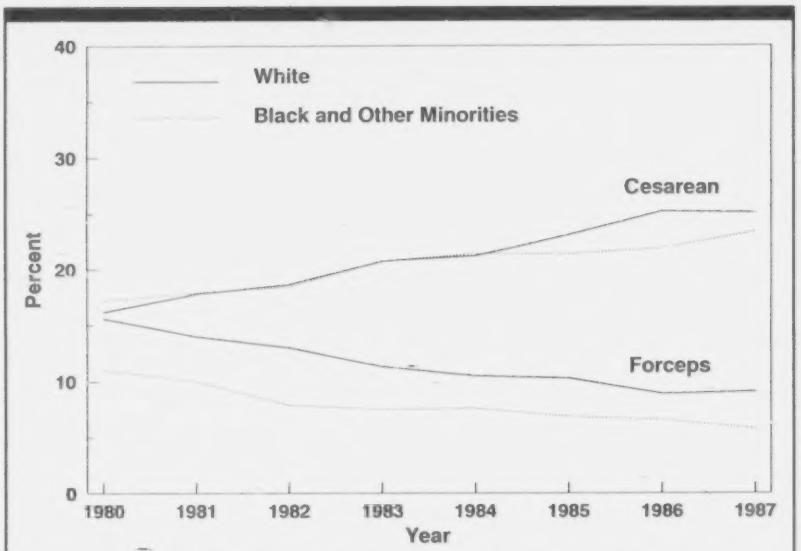


FIGURE 2—Rate of US cesarean sections and forceps procedures, by race, 1980 to 1987, National Hospital Discharge Survey, National Center for Health Statistics, Hyattsville, Md.

Length of Stay

Despite the increase in the percentage of all deliveries associated with an operative procedure, the mean length of hospital stay for all deliveries decreased from 3.8 days in 1980 to 3.1 days in 1987 (test for linear trend, $P = .0001$). During the same period, the mean length of hospital stay for cesarean section deliveries declined from 6.5 to 4.8 days (test for linear trend, $P = .0001$); for forceps deliveries, it decreased from 3.5 to 2.8 days (test for linear trend, $P = .0001$); and for vacuum extractions, it declined from 3.2 to 2.7 days (test for linear trend, $P = .002$).

Age

The risk of having a cesarean section increased with the increasing age of the mother. Women 35 years of age and older were 30% more likely to receive a cesarean section than were women 20 through 34 years of age (Table 1). In contrast, there were no significant differences in the rates of forceps or vacuum extraction procedures among women of different ages. From 1980 to 1987, the rate of cesarean sections increased significantly in each age group (test for linear trend, $P < .002$ in each age group).

Race

Since 1980, rates for cesarean sections have increased 55% among Whites and 36% among Black and other minority women (test for linear trend, $P = .0001$ for each racial group) (Figure 2). Forceps procedures declined 41% among White women and 46% among Black and other minority women (test for linear trend, $P = .0001$ for White women and $P = .0008$ for Black and other minority women). In 1987, White women were 60% more likely to have a forceps delivery than were Black and other minority women (Table 1).

Insurance Status

For each year in the study period, insured women had higher rates of cesarean sections, forceps procedures, and vacuum extraction procedures than did uninsured women. In 1987, women with private insurance were 20% more likely to have a cesarean section, 70% more likely to have a delivery by forceps, and 80% more likely to have a vacuum extraction procedure than women without private insurance (Table 1). To further investigate the relationship between operative delivery procedures and race, we stratified by insurance status (Table 2). Within each insurance category, the rate of each procedure was not significantly different between the two racial groups.

Geographic Region

The rate of specific intervention procedures differed slightly by geographic region within the United States (Table 1). For example, women in the South were more than twice as likely to have a forceps delivery as women in the West, and women in the Northeast were 20% more likely to be delivered by cesarean section than were women in the West. Although women in the West had the highest rate of vacuum extraction procedures, this rate was not statistically different from that of other regions.

Discussion

In 1987, an operative procedure was performed on more than a third of all women in delivery. Although the rate of cesarean sections increased by 48% between 1980 and 1987, the rate of forceps

procedures declined by 43%. During this same period, Kozak reported a 42% increase in the rates of a variety of surgical and nonsurgical procedures associated with US hospital deliveries.[9] Cesarean sections, repair of obstetric lacerations, artificial rupture of membranes, and vacuum extraction accounted for more than half the increased number of delivery procedures. Of the 11 procedures Kozak evaluated, only the use of forceps declined.

Our analysis of procedure rates by maternal age showed that older women were significantly more likely to receive a cesarean section than were younger women—a finding supported by previous studies,[2,10,11] even when researchers controlled for other risk factors such as preexisting hypertension and diabetes.[12] In contrast, the rates of forceps and vacuum extraction procedures did not increase among older women. Obstetricians may apply different criteria when choosing a method of delivery for older women, whom they consider to be at high risk.[12] This may explain why the rate of forceps procedures was slightly lower among women 35 years of age and older in our study.

Disparities in access to and receipt of health care services for Blacks have been previously reported.[13] Other studies have looked at the difference in cesarean section rates by race and ethnicity, often in combination with socioeconomic status or payment source.[14,15] Gould et al. controlled for family income and found no significant difference in cesarean rates between non-Hispanic White and Black women.[15] They did find significantly higher rates among women of higher income levels regardless of race, however. We found that White women had the same risk of receiving a cesarean section as Black and other minority women, while the risk of forceps procedures was significantly higher for White women. When we stratified by insurance status, the racial difference was no longer seen for any of the delivery procedures.

Private patients and those with private insurance tend to have the highest rates of cesarean sections, whereas clinic and uninsured patients tend to have the lowest rates, regardless of risk status.[14,16–19] Women who obtain services from a health maintenance organization tend to have intermediate cesarean rates.[14] Other surgical procedures, including hysterectomy and tonsillectomy, also are performed more frequently when the physician is reimbursed on a fee-for-service basis.[20] Our data indicate that cesarean sections are 20% more likely to occur among insured women than they are among women without private insurance. Moreover, forceps and vacuum extraction procedures are 70% to 80% more likely to be performed among women with private insurance than among those without private insurance.

Geographic variations in health care practices[19,21–23] and higher cesarean section rates in the Northeast and South have been previously reported, but information on the geographic variation in forceps and vacuum extraction procedures has not previously been published. We found the rate of forceps procedures to be significantly greater in the South.

Some researchers have suggested that the increase in cesarean sections is partly responsible for the decline in forceps procedures.[2] Our findings are consistent with this hypothesis, although we cannot test its validity with the National Hospital Discharge Survey data. Information about additional factors that can influence the rates of cesarean and forceps procedures (e.g., parity, fetal conditions during labor, birthweight) is not available within this data set. In the near future, however, additional information will be available to examine indications and outcomes of operative delivery procedures. The revised US Standard Certificate of Live Birth was implemented by all states in 1991, and it contains a question on method of delivery.[24] This will facilitate the evaluation of operative delivery procedures by birthweight, gestational age, Apgar score, and many abnormal conditions of the newborn.

The Healthy People 2000 Objectives call for decreasing the national cesarean section rate from 24% to no more than 15% of all deliveries (the objectives do not address other operative delivery procedures).[25] Numerous strategies have been proposed to lower the cesarean section rate in the United States.[23,26] For example, in two programs that reduced cesarean section rates to about 11% (from 18% and 28%),[27,28] a comprehensive system of peer review was established, and all cesarean sections were discussed at staff conferences. Standardized protocols and required second opinions have also been found to be effective.[23,27,28] The protocol of Myers and Gleicher led to a decline in the rate of cesarean deliveries (from 17.5% in 1985 to 11.5% in 1987) and in the rate of operative vaginal deliveries (from 10.4% in 1985 to 4.3% in 1987).[27] Sanchez-Ramos and colleagues reported that their rate of operative vaginal deliveries remained stable as the rate of cesarean sections decreased (from 27.5% in 1986 to 10.5% in 1989).[28] In 1989, however, their rate of operative vaginal procedures was 18.5%, far greater than the national rate.

As the pressure mounts to reduce the number of cesarean sections performed in US hospitals, we also should focus our attention on the surveillance of other operative delivery procedures. Prospective studies of maternal and infant characteristics and outcomes are needed to help us determine which procedures can produce the best outcomes for mothers and infants. □

Acknowledgments

An earlier version of this paper was presented at the 118th Annual Meeting of the American Public Health Association, October 1990, New York City.

We are grateful to Lola Jean Kozak, RN, PhD, Division of Health Care Statistics, National Center for Health Statistics, for her review of this manuscript.

References

1. Taffel SM, Placek PJ, Moien M. Cesarean section rate levels off in 1987. *Fam Plann Perspect.* 1989;21:227–228.
2. Taffel SM, Placek PJ, Liss T. Trends in the United States cesarean–section rate and reasons for the 1980–85 rise. *Am J Public Health.* 1987;77:955–959.
3. Notzon FC. International differences in the use of obstetric interventions. *JAMA.* 1990;263:3286–3291.
4. Graves EJ. Detailed diagnoses and procedures, National Hospital Discharge Survey: 1987. *Vital Health Stat [13].* 1989;100. DHHS publication PHS 89-1761.
5. *International Classification of Diseases, 9th Revision, Clinical Modification.* Washington, DC: US Department of Health and Human Services; 1991. DHHS publication PHS 91-1260.
6. Shah BV. *SESUDAAN: Standard Errors Program for Computing of Standardized Rates from Sample Survey Data.* Research Triangle Park, NC: Research Triangle Institute; 1981.
7. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research.* New York, NY: Van Nostrand Reinhold; 1982:297–299.
8. Neter J, Wasserman W, Kutner MH. *Applied Linear Regression Models.* 2nd ed. Boston, Mass: Richard D Irwin; 1989:418–424.
9. Kozak LJ. Surgical and nonsurgical procedures associated with hospital delivery in the United States: 1980–1987. *Birth.* 1989; 16:209–213.
10. Berkowitz GS, Fiarman GS, Mojica MA, Bauman J, Haynes de Reg: R. Effect of physician characteristics on the cesarean birth rate. *Am J Obstet Gynecol.* 1989;161: 146–149.
11. Gordon D, Milberg J, Daling J, Hickok D.

Advanced maternal age as a risk factor for cesarean delivery. *Obstet Gynecol.* 1991; 77:493–497.
12. Grimes DA, Gross GK. Pregnancy outcomes in Black women aged 35 and older. *Obstet Gynecol.* 1981;58:614–620.
13. Council on Ethical and Judicial Affairs. Black-White disparities in health care. *JAMA.* 1990;263:2344–2346.
14. Kizer K, Ellis A. C-section rate related to payment source. *Am J Public Health.* 1988;78:96–97.
15. Gould JB, Davey B, Stafford RS. Socioeconomic differences in rates of cesarean section. *N Engl J Med.* 1989;321:233–239.
16. Haynes de Regt R, Minkoff HL, Feldman J, Schwarz RH. Relation of private or clinic care to the cesarean birth rate. *N Engl J Med.* 1986;315:619–624.
17. Wilner S, Schoenbaum SC, Monson RR, Winickorr RN. A comparison of the quality of maternity care between a health-maintenance organization and fee-for-service practices. *N Engl J Med.* 1981;304:784–787.
18. Stafford RS. Cesarean section use and source of payment: an analysis of California hospital discharge abstracts. *Am J Public Health.* 1990;80:313–315.
19. Placek PJ, Taffel SM. Recent patterns in cesarean delivery in the United States. *Obstet Gynecol Clin North Am.* 1988;15:607–627.
20. LoGerfo JP, Efird RA, Diehr PK, Richardson WC. Rates of surgical care in prepaid group practices and the independent setting: what are the reasons for the differences? *Med Care.* 1979;17:1–7.
21. Berkowitz GS, Skovron ML, Lapinski RH, Berkowitz RL. Delayed childbearing and the outcome of pregnancy. *N Engl J Med.* 1990;322:659–664.
22. Goyert GL, Bottoms SF, Treadwell MC, Nehra PC. The physician factor in cesarean birth rates. *N Engl J Med.* 1989;320:706–709.
23. Shiono PH, Fielden JG, McNellis D, Rhoads GG, Pearse WH. Recent trends in cesarean birth and trial of labor rates in the United States. *JAMA.* 1987;257:494–497.
24. Taffel SM, Placek PJ, Moien M, Kosary CL. 1989 U.S. cesarean section rate steadies—VBAC rate rises to nearly one in five. *Birth.* 1991;18:73–77.
25. *Healthy People 2000: National Health Promotion and Disease Prevention Objectives.* Washington, DC: US Department of Health and Human Services, Public Health Service; 1990. DHHS publication PHS 91-50212.
26. Stafford RS. Alternative strategies for controlling rising cesarean section rates. *JAMA.* 1990;263:683–687.
27. Myers SA, Gleicher N. A successful program to lower cesarean-section rates. *N Engl J Med.* 1988;319:1511–1516.
28. Sanchez-Ramos L, Kaunitz AM, Peterson HB, Martinez-Schnell B, Thompson RJ. Reducing cesarean sections at a teaching hospital. *Am J Obstet Gynecol.* 1990;163: 1081–1087.

Changes in Mammography Use: Economic, Need, and Service Factors

ABSTRACT

Objectives. **The purpose of this study was to document changes in mammography use between 1987 and 1990 and assess the relationship of use to selected economic, need, and health system factors.**

Methods. **Independent random-digit-dialed telephone surveys of women between 52 and 75 years of age were conducted.**

Results. **Between 1987 and 1990, the proportion of women who had had a mammogram in the past year increased from 31% to 51%. Though income was significantly related to overall patterns of use, it was not associated with recent mammogram use in 1990. Women with a family history of breast cancer reported greater use at both times, as did women who reported having a regular physician (particularly a gynecologist or internist). When all other variables were controlled for, women were over nine times more likely to have had multiple and recent mammograms in 1990 than in 1987.**

Conclusions. **Mammography use dramatically increased between 1987 and 1990. There were strong relationships between the type of regular physician and mammography screening and between economic and personal history and repeated and recent mammography use. (*Am J Public Health.* 1992;82:1345–1351)**

Jane G. Zapka, ScD, David Hosmer, PhD, Mary E. Costanza, MD, Donald R. Harris, PhD, and Anne Stoddard, ScD

Introduction

Breast cancer continues to be a prominent women's health problem, representing 28% of all female cancers.[1] Since reductions in mortality depend on increasing the adoption of early detection methods by physicians and by women, numerous professional organizations have endorsed screening guidelines for periodic clinical breast examinations and mammography.[2] If the Year 2000 Health Goals for the Nation are to be achieved, a better understanding of factors related to screening services use is needed.[3]

In this study we document changes in mammography utilization over a 3-year period and investigate the relationship of selected economic, need, and health system factors to mammography screening utilization to determine which subgroups of women are being screened.

Methods

Study Subjects and Data Collection

Independent random-digit-dialed surveys of English-speaking women between 52 and 75 years of age who had never had breast cancer were conducted in two Massachusetts urban areas comprising six towns. The surveys were conducted in the fall of 1987 (n = 929) and the fall of 1990 (n = 674) as part of a 3-year demonstration project, sponsored by the National Cancer Institute (NCI), designed to increase breast cancer screening.[4] The study sample was drawn from the demonstration and comparison urban areas. The interviews, averaging 17 minutes in length, were conducted during day and evening hours, with a minimum of 10 call-backs at varying times. A refusal conversion call procedure was used, and response rates of 75.3% in 1987 and 73.9% in 1990 were achieved. The 52-year age minimum was based on professional guidelines recommending annual screening for women aged 50 years and older. Therefore, a 52-year-old woman could have adhered to the recommendation at least twice.

Study Variables

The dependent variable, mammography utilization, was defined with four categories: (1) no mammogram (never had one); (2) previous user (had had one or more mammograms, but none in the previous 12 months); (3) recent adopter (had had only one mammogram, which had occurred in the previous 12 months); and (4) repeat recent user (had had two or more mammograms, including one in the previous 12 months). This variable was designed to provide a better understanding of factors influencing periodic as well as recent utilization. For a subsequent analysis, the categories were collapsed (1 plus 2; 3 plus 4) to form a binary dependent variable, defined as having had a mammogram in the previous 12 months or not. Because of missing data, the 1987 and 1990 samples consisted of 838 and 601 women, respectively.

In the selection of study variables, which was guided by previous work,[5–7] we focused on three groups of factors

Jane G. Zapka, David Hosmer, Donald R. Harris, and Anne Stoddard are with the School of Public Health, University of Massachusetts, Amherst. Mary E. Costanza is with the University of Massachusetts Medical School, Worcester, Mass.

Requests for reprints should be sent to Jane G. Zapka, ScD, School of Public Health, University of Massachusetts, Amherst, MA 01003.

This paper was submitted to the Journal December 30, 1991, and accepted with revisions May 5, 1992.

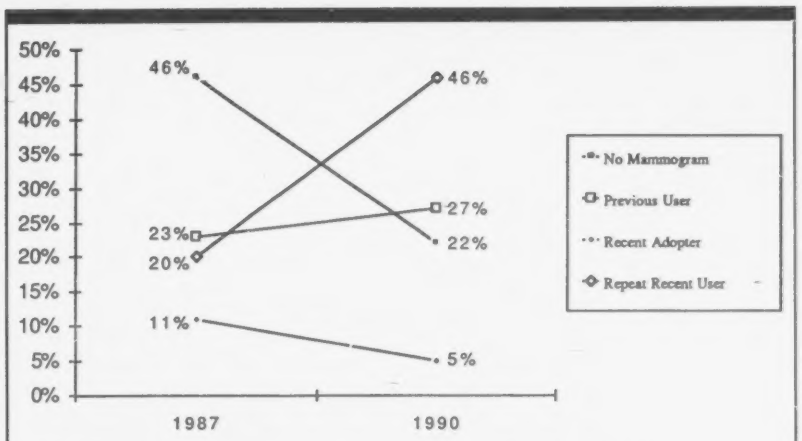


FIGURE 1—Percentages of women using mammograms, by pattern of use.

(economic, need, and health system enabling factors) and their relationships to screening. Economic factors included measures of income, education, and employment. Personal health need factors, such as family history of breast cancer, personal history of breast symptoms, and age, may influence use of screening by influencing women's personal perception of risk, their physicians' attention to screening recommendations, or both. Access to and interaction with health providers may facilitate or hinder mammography use. A physician's advice that a client obtain a mammogram was one health system enabling factor investigated in this study; others included having insurance coverage, having a regular physician, and having had a clinical breast exam in the past year.

Analysis

A univariate examination of the relationship between study variables and mammography use was performed for each survey. To better control for potential confounding, variables significant at the .15 level in either survey were selected as candidates for multivariate modeling. A polytomous logistic regression model was used to assess the joint effects of the variables.[8] Likelihood ratio tests reduced the model to that subset of variables significant at the .10 alpha level in either survey. An indicator variable for city was included in all models to control for intervention effect. To check for possible shifts in effects between survey waves, a combined model was fit, and interactions between time and study variables were tested by means of likelihood ratio tests. None were significant at the 5% level. The adequacy of the polytomous logistic regression model was assessed by applying logistic regression diagnostics to equivalent separate sample binary logistic regression models. In the second phase of the analysis, models employing the binary dependent variable (mammogram in the past 12 months or not) were fit to the data.

Results

Mammography Utilization

The pattern of mammography use was significantly related to study period ($\chi^2 = 151.2$; $P < .0001$, $df = 3$). The proportion of women who had had a mammogram in the past 12 months (the recent adopters plus the repeat recent users) increased from 31% in 1987 to 51% in 1990 (Figure 1); a corresponding decrease occurred in the proportion of women who had never had a mammogram, from 46% to 22%. Both shifts were statistically significant ($P < .0001$). The proportion of previous users increased from 23% to 27%; the proportion of recent adopters decreased from 11% to 5%; and the proportion of repeat recent users increased from 20% to 46%. Only the latter two shifts were statistically significant ($P \leq .0002$).

Relationship of Study Variables to Mammography Use

Economic variables. The relationship of employment status to screening use was of borderline significance in 1987, but not in 1990 (Table 1). In 1987, unemployed women were somewhat more likely to have never had a mammogram, whereas employed women were slightly more likely to be previous users. Education and income level were significantly related to utilization at both times. Women with less than a high school education were most likely to have never had a mammogram, whereas college graduates were most likely to report the highest utilization in both surveys. Women with total household incomes of $15 000 or less were most likely to have never had a mammogram and least likely to be previous users or recent users. More women in the lower income groups reported having had a recent mammogram in 1990 than in 1987, but their rates were still lower than those of high-income women.

Need variables. Women with a family history of breast cancer were significantly more likely to have had screening in 1987, but not in 1990 (Table 1). Women with a personal history of breast symptoms or problems were significantly more likely to be previous users and repeat recent users in both surveys. In 1987, women 65 years of age and older were more likely to have never had a mammogram, but were about as likely to have had one in the past year as younger women. By 1990, the two age groups had very similar patterns of use.

Health system enabling variables. In both 1987 and 1990, having either an internist or a gynecologist as a regular physician was related to ever having had a mammogram, as well as to being a repeat recent user (Table 1). In both surveys, women who reported being advised by a physician to have a mammogram had higher levels of use. Having had a recent clinical breast examination was significantly and consistently associated with ever having had a mammogram and with being a repeat recent user.

Insurance coverage was significantly related to mammography utilization at both times (Table 1). However, analyses of the relationship between insurance coverage and screening utilization should be interpreted with caution. Because some women had multiple sources of insurance, it was not possible to determine which plan actually provided coverage. In addition, coverage may change annually, so comparing mammography use over time with current coverage is of dubious validity. Therefore, we investigated the relationship of insurance coverage to mammogram use in the previous year (recent adopters and repeat recent users combined). Though this relationship was of borderline significance in 1987 and nonsignificant in 1990, certain trends are noteworthy (Figure 2). Women with no insurance were least likely to have had a mammogram in the past year in both 1987

TABLE 1—Percentage Distribution of Mammography Use for Study Variables for 1987 and 1990

	1987 (n = 838)						1990 (n = 601)					
	No.	No Mammogram, % (n = 385)	Previous User, % (n = 195)	Recent Adopter, % (n = 92)	Repeat Recent User, % (n = 166)	χ^2 P Value	No.	No Mammogram, % (n = 130)	Previous User, % (n = 162)	Recent Adopter, % (n = 32)	Repeat Recent User, % (n = 277)	χ^2 P Value
Economic factors												
Employed												
Yes	350	41.4	27.4	11.7	19.4		239	19.3	29.7	4.2	46.9	
No	488	49.2	20.3	10.5	20.1	.058	362	23.2	25.1	6.1	45.6	.365
Education												
< High school	172	59.9	16.9	10.5	12.8		142	31.0	30.3	8.5	30.3	
High school graduate	383	44.4	26.4	9.1	20.1		247	20.7	27.1	5.7	46.6	
Some college/ technical school	183	41.0	24.0	15.9	19.1		126	16.7	24.6	2.4	56.4	
College graduate +	100	37.0	21.0	10.0	32.0	<.001	84	16.7	23.8	3.6	56.0	.001
Income												
≤ $15 000	286	58.4	18.2	12.2	11.2		177	30.5	23.2	9.6	36.7	
$15 001–25 000	155	40.7	29.7	11.6	18.1		116	20.7	30.2	8.6	40.5	
$25 001–35 000	110	41.8	26.4	7.3	24.6		75	20.0	28.0	1.3	50.7	
>$35 000	165	33.3	24.2	12.7	29.7	<.001	162	16.1	32.1	2.5	49.4	.002
Need factors												
Family history of breast cancer												
Yes	92	30.4	25.0	14.1	30.4		69	18.8	20.3	7.3	53.6	
No	746	47.9	23.1	10.6	18.5	.006	532	22.0	27.8	5.1	45.1	.382
Prior breast problems												
Yes	193	18.1	38.9	7.8	35.2		167	3.6	30.5	1.2	64.7	
No	645	54.3	18.6	11.9	15.2	<.001	433	28.6	25.6	6.9	38.8	<.001
Age group												
≤ 64 y	529	41.2	26.7	9.8	22.3		374	20.3	28.3	5.1	46.3	
≥ 65 y	309	54.1	17.5	12.9	15.5	<.001	227	23.8	24.7	5.7	45.8	.657
Health system enabling factors												
Specialty of regular physician												
None	124	66.1	23.4	3.2	7.3		76	48.7	25.0	1.3	25.0	
General or family practice	242	61.2	16.5	10.7	11.6		195	26.7	31.3	6.7	35.4	
Internist	264	34.9	26.5	12.1	26.5		217	11.1	27.2	6.9	54.8	
Gynecologist	161	27.3	26.1	14.9	31.7		84	3.6	22.6	2.4	71.4	
Other	47	40.4	29.8	12.8	17.0	<.001	29	48.3	13.8	3.5	34.5	<.001
Physician advised patient to get mammogram												
Yes	468	9.4	38.9	17.7	34.0		472	6.8	30.3	6.6	56.4	
No	331	91.2	3.9	2.7	2.1	<.001	111	73.9	16.2	0.9	9.0	<.001
Last clinical breast exam												
≤ 12 mo ago	493	33.3	18.1	17.2	31.4		363	10.5	17.6	6.3	65.6	
Never/> 12 mo ago	344	64.2	30.8	2.0	2.9	<.001	238	38.7	41.2	3.8	16.4	<.001
Insurance												
None	24	83.3	8.3	8.3	0.0		15	66.7	6.7	0.0	26.7	
Entitlement only	54	64.8	11.1	9.3	14.8		90	26.7	23.3	10.0	40.0	
Entitlement +	260	50.8	20.0	12.7	16.5		123	26.0	21.1	5.7	47.2	
Health maintenance organization only	52	28.9	30.8	7.7	32.7		74	13.5	27.0	1.4	58.1	
Health maintenance organization + commercial	42	42.9	21.4	11.9	23.8		22	9.1	36.4	9.1	45.5	
Commercial only	358	40.5	26.0	10.9	22.6	<.001	204	17.2	30.4	5.9	46.6	.001

and 1990. Women with entitlements only or entitlements plus other coverage did substantially better in 1990 than in 1987. In both years, women with health maintenance organization coverage were more likely than women in any other coverage group to have had a mammogram in the past year.

Aside from the significant relationships of the independent variables with utilization at each time point, notable changes occurred over time. For example, more women were advised by their physicians to have a mammogram in 1990 than in 1987 (81% vs 59%). Although the proportion of women who had had a clinical breast exam in the previous year did not change (59% in 1987 and 60% in 1990),

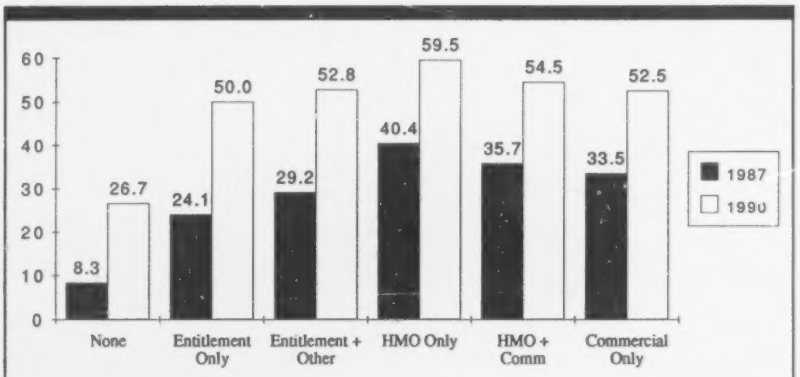


FIGURE 2—Percentages of women having a mammogram in the previous 12 months, by insurance coverage.

more women who had had such an exam reported having had a mammogram in 1990 (72%) than in 1987 (49%).

Multivariate Analyses

Multivariate analysis was used to investigate the relationship of specific variables to utilization while controlling for the effects of others. The variable "physician advised patient to get a mammogram" was not included because of its high correlation with the dependent variable. Insurance coverage was omitted because of its lack of validity with respect to utilization in other than the past year. The fitted logistic regression models are presented in Tables 2 and 3 in the form of odds ratios, obtained by exponentiating the estimated coefficients, and confidence intervals for the odds ratios.

In Table 2, odds ratios are given that compare the referent group, women who had never had a mammogram, with the three user groups: previous users, recent adopters, and repeat recent users. Consequently, for each dichotomous variable in Table 2, there are three odds ratios. For the variable "specialty of regular physician," four sets of odds ratios are given, one for each physician specialty compared with having no regular physician. The income variable was treated as continuous and the odds ratio for an increase of $5000 is presented.

The first column of odds ratios in Table 2 reports the results of the model for the 1987 survey. When other important variables were controlled for, economic factors retained their significance. Employed women were less likely than unemployed women to be repeat recent users. Women with at least a high school education were more likely than women with less than a high school education to be previous users. Higher income increased the likelihood of a woman's being a repeat recent user.

With regard to need factors, women who reported a family history of breast cancer were approximately three times more likely than others to be recent adopters or repeat recent users in 1987. Women who had had breast problems were more than six times as likely as others to have ever had more than one mammogram. The likelihood that a woman who had had symptoms was a recent adopter, though not significant, was nearly twice as high as the likelihood that she had never had a mammogram. Women younger than 65 years of age were twice as likely as older women to be repeat recent users.

With regard to health system enabling factors, women whose regular physician was an internist or gynecologist reported greater use than did women with no regular physician. In addition, they were almost eight times more likely than women who regularly visited a general practitioner or family physician to be repeat recent users. If a woman had had a clinical breast exam in the past 12 months, she was 12 to 18 times more likely to have had a recent mammogram. Women in the demonstration city area were less likely to report previous use than were those in the comparison city area; recent use was similar.

The same model fit to the 1990 survey data is displayed in the center columns of Table 2. Variables significant in this model included education, previous breast problems, specialty of regular physician, and recency of clinical breast exam. The overall relationships between mammography use and having had a breast problem and recency of clinical breast exam were the same as in 1987, although the odds ratios changed somewhat. Women were far more likely to be recent adopters if they had a regular physician (21 to 22 times more likely if the physician was a gynecologist or internist, and almost 9 times more likely if the physician was a general practitioner or family physician). Women were also more likely to be repeat recent users if they reported having a regular physician (from 2 to 17 times more likely, depending on the specialty of the regular physician).

In the next stage of analysis (final columns of Table 2), we fit the same model to the combined survey data, including an indicator variable for time. In the combined model, when other significant variables were controlled for, two economic factors achieved statistical significance. Women with more than a high school education were more than twice as likely to be repeat recent users, and a $5000 increase in income increased the likelihood that a woman would be a repeat recent user. With respect to need factors, women with a family history of breast cancer were twice as likely as other women to be repeat recent users; those with a history of breast problems were more than eight times as likely to be repeat recent users. Age was not significantly related to use when the influence of all other variables was controlled. Among the health system enabling factors, specialty of regular physician continued to be highly significant. Women whose regular physicians were gynecologists or internists were more likely to have ever had a mammogram. When all other variables were controlled for, women were more than nine times as likely to be repeat recent users in 1990 than in 1987.

Because several national reports have focused on the number of women who have had mammograms in the 12 months prior to study,[4,9] models were fit to the dichotomous outcome variable (having had a mammogram in the past 12 months or not). The models for 1987, for 1990, and combined over time are presented in Table 3. Economic variables were not significant, except for income in 1987 and education in 1990. Having had previous breast problems was related to use in both 1987 and 1990, whereas having a family history of breast cancer was related in 1987 only; age was not related. Having as a regular physician a gynecologist (or an internist in the 1990 and combined-time models) and having had a clinical breast exam in the past 12 months were related to increased likelihood of recent use. A model that included insurance rather than income yielded similar results.

TABLE 2—Estimated Odds Ratios and 95% Confidence Intervals (CI) for Mammography Utilization in 1987, in 1990, and Combined over Time

Effect	Mammography Utilization (vs None)	1987		1990		Combined	
		Odds Ratio	95% CI	Odds Ratio	95% CI	Odds Ratio	95% CI
Economic factors							
Employed	Previous user	0.95	0.59, 1.52	1.20	0.63, 2.29	1.03	0.71, 1.49
(Yes vs no)	Recent adopter	1.07	0.58, 1.99	1.20	0.41, 3.49	1.09	0.65, 1.84
	Repeat recent user	0.51	0.29, 0.90	1.11	0.56, 2.21	0.70	0.46, 1.07
Education	Previous user	1.92	1.04, 3.55	1.09	0.56, 2.12	1.33	0.87, 2.04
(High school + vs < high school)	Recent adopter	0.73	0.36, 1.46	1.31	0.50, 3.40	0.87	0.50, 1.51
	Repeat recent user	1.47	0.69, 3.16	2.89	1.42, 5.89	2.09	1.27, 3.45
Income	Previous user	1.05	0.97, 1.13	1.09	0.99, 1.20	1.08	1.02, 1.14
($5000	Recent adopter	1.08	0.98, 1.19	0.97	0.72, 1.03	1.02	0.93, 1.11
increments)	Repeat recent user	1.17	1.07, 1.28	1.03	0.93, 1.14	1.12	1.04, 1.20
Need factors							
Family history	Previous user	1.83	0.85, 3.93	0.63	0.26, 1.54	1.23	0.69, 2.19
(Yes vs no)	Recent adopter	2.76	1.17, 6.54	0.76	0.21, 2.78	1.90	0.95, 3.81
	Repeat recent user	3.53	1.59, 7.85	0.88	0.37, 2.08	2.00	1.12, 3.58
Prior breast problems	Previous user	6.52	3.86, 11.03	9.31	3.65, 23.72	7.39	4.72, 11.56
(Yes vs no)	Recent adopter	1.81	0.85, 3.85	1.19	0.22, 6.46	1.67	0.85, 3.29
	Repeat recent user	7.73	4.23, 14.14	10.99	4.27, 28.24	8.63	5.34, 13.96
Age group	Previous user	1.62	0.96, 2.74	1.15	0.59, 2.24	1.39	0.93, 2.08
(<65 y vs ≥65 y)	Recent adopter	0.96	0.50, 1.85	0.95	0.35, 2.57	0.94	0.55, 1.60
	Repeat recent user	1.93	1.03, 3.60	0.84	0.43, 1.64	1.25	0.80, 1.94
Health system enabling factors							
Specialty of regular physician	Previous user	0.65	0.33, 1.30	2.19	1.05, 4.56	1.15	0.70, 1.88
(General/family practice vs	Recent adopter	1.42	0.44, 4.64	8.77	1.07, 71.78	2.72	1.00, 7.38
none)	Repeat recent user	0.88	0.29, 2.62	2.19	0.96, 4.98	1.43	0.75, 2.71
(Internist vs none)	Previous user	2.32	1.20, 4.49	5.15	2.24, 11.83	3.19	1.91, 5.33
	Recent adopter	2.60	0.79, 8.48	20.72	2.45, 175.15	5.18	1.89, 14.19
	Repeat recent user	3.01	1.05, 8.65	7.00	2.92, 16.76	4.37	2.30, 8.30
(Gynecologist vs none)	Previous user	3.27	1.53, 7.00	9.98	2.46, 40.39	4.66	2.48, 8.76
	Recent adopter	4.60	1.32, 16.05	22.22	1.46, 338.86	7.78	2.58, 23.43
	Repeat recent user	6.72	2.21, 20.40	16.91	4.07, 70.27	9.26	4.40, 19.51
(Others vs none)	Previous user	1.77	0.64, 4.91	0.42	0.11, 1.67	1.03	0.46, 2.31
	Recent adopter	1.17	0.24, 5.78	1.70	0.09, 30.89	1.50	0.39, 5.81
	Repeat recent user	1.43	0.35, 5.78	0.74	0.21, 2.59	1.21	0.48, 3.03
Last clinical breast exam	Previous user	1.07	0.68, 1.69	1.16	0.64, 2.08	1.10	0.78, 1.56
(≤ 12 mo vs	Recent adopter	12.36	5.35, 28.57	5.12	2.04, 12.86	7.86	4.38, 14.09
none or earlier)	Repeat recent user	17.76	7.95, 39.71	9.35	5.11, 17.10	10.51	6.80, 16.24
Other factors							
City	Previous user	0.64	0.42, 0.98	0.59	0.34, 1.04	0.64	0.46, 0.89
(Demonstration vs	Recent adopter	1.29	0.75, 2.22	1.10	0.46, 2.63	1.24	0.79, 1.95
comparison)	Repeat recent user	0.63	0.38, 1.05	0.98	0.55, 1.74	0.79	0.55, 1.14
Time	Previous user					3.52	2.47, 5.02
(1990 vs 1987)	Recent adopter					1.65	0.99, 2.77
	Repeat recent user					9.09	6.11, 13.51

Note. Likelihood ratio test of significance of 1987 model: 372.8 ($df = 36$); $P < .001$. Deviance = 1364.6 ($df = 1953$).
Likelihood ratio test of significance of 1990 model: 278.5 ($df = 36$); $P < .001$. Deviance = 982.4 ($df = 1437$).
Likelihood ratio test of significance of combined model: 713.5 ($df = 39$); $P < .001$. Deviance = 2400.3 ($df = 3489$).

Discussion

The results of this study support recently published data that document improvements in the proportion of women who have had a mammogram in the year prior to study.[9] In addition, this study documents an increase in the proportion of women who have been screened recently among those who have had mammograms in the past (20% in 1987 to 46% in 1990). The challenge will be to promote regular screening by new adopters, to continue to encourage the substantial proportion (approximately 25%; see Figure 1) of previous (but not recent) users to again participate in screening, and to motivate the hard-to-reach 22% who have never been screened. One limitation of this study is the lack of representation of minority and non-English-speaking women; more than 90% of each sample was White. The study is also limited by the usual biases of random-digit-dialed surveys: only women with telephones could participate and the data were self-reported.

Among the economic factors, income was significantly related to repeat recent

TABLE 3—Estimated Odds Ratios and 95% Confidence Intervals for Recent Users of Mammography with Past or Never Users in 1987, in 1990, and Combined over Time

Effect	1987		1990		Combined	
	Odds Ratio	95% CI	Odds Ratio	95% CI	Odds Ratio	95% CI
Economic factors						
Employed (Yes vs no)	0.71	0.46, 1.10	0.97	0.59, 1.61	0.78	0.57, 1.09
Education (High school+ vs < High school)	0.84	0.48, 1.46	2.27	1.32, 3.90	1.34	0.92, 1.92
Income ($5000 increments)	1.11	1.04, 1.19	0.96	0.88, 1.03	1.04	0.99, 1.09
Need factors						
Family history (Yes vs no)	2.36	1.29, 4.32	1.16	0.61, 2.21	1.75	1.13, 2.70
Prior breast problems (Yes vs no)	1.77	1.15, 2.73	1.69	1.06, 2.69	1.71	1.25, 2.34
Age group (<65 y vs ≥65 y)	1.18	0.73, 1.89	0.81	0.49, 1.35	0.98	0.70, 1.38
Health system enabling factors						
Specialty of regular physician						
(General/family practice vs none)	1.24	0.54, 2.87	1.80	0.90, 3.63	1.59	0.93, 2.70
(Internist vs none)	2.00	0.88, 4.55	3.02	1.52, 6.02	2.48	1.47, 4.19
(Gynecologist vs none)	3.42	1.45, 8.06	3.89	1.67, 9.06	3.82	2.14, 6.83
(Others vs none)	1.09	0.37, 3.20	1.11	0.38, 3.24	1.22	0.50, 2.56
Last clinical breast exam (≤ 12 mo vs none or earlier)	14.73	8.23, 26.36	7.48	4.82, 11.61	9.06	6.49, 12.65
Other factors						
City (Demonstration vs comparison)	1.03	0.70, 1.51	1.39	0.91, 2.12	1.13	0.85, 1.49
Time (1990 vs 1987)					2.91	2.17, 3.89

Note. Likelihood ratio test of significance of 1987 model: 215.1 ($df = 12$); $P < .001$. Deviance = 634.0 ($df = 651$).
Likelihood ratio test of significance of 1990 model: 183.0 ($df = 12$); $P < .001$. Deviance = 550.9 ($df = 479$).
Likelihood ratio test of significance of combined model: 367.7 ($df = 13$); $P < .001$. Deviance = 1207.5 ($df = 1163$).

mammography use in 1987 and in the combined model (Table 2); however, income was not significantly related to recent mammography use in 1990 (Table 3). This change may be due, in part, to the fact that more women with entitlements were being screened in 1990 (Figure 2).

With respect to need factors, having had breast problems was significant in all three models, whereas having a family history of breast cancer was significant only in the 1987 and combined models. Both factors were also significantly related to recent mammography use in 1987 and in the combined model (Table 3). It could be hypothesized that clinicians in 1987 were more likely to advise screening for these women because of their health history, but more recently, all women have been advised to have regular mammograms. It is not surprising that a history of breast symptoms or problems is independently related to mammography use. In both 1987 and 1990, 17% of the women reported that they had had their most recent mammogram because they had had a breast problem, rather than as part of a regular checkup. It may be that these women misunderstood why the mammogram was recommended. For example, considerable anecdotal evidence suggests that mammograms are ordered to rule out problems in order to secure insurance coverage.

The bivariate relationship of age and family history to increased use in 1987 and 1990, coupled with increases in use over time, supports the notion that screening of all women is increasingly being encouraged. The relationship of income and education to use supports the recommendation that special attention be given to the underserved, that is, women with lower incomes and less education. In this regard, it is encouraging that in 1990 more women with entitlements only and more uninsured women were receiving care. However, the fact that more elderly women were being screened provides additional evidence that factors other than income influence mammography use.[10-12]

With respect to health system enabling factors, this investigation demonstrates the continued profound relationships between having a regular physician and mammography use and between physician specialty and mammography use that have been documented in earlier studies.[4,13-15] Although having a gynecologist as a regular physician continues to be important, women reporting internists as regular physicians have improved their use of mammography in recent years. Since women who indicated that their regular physician was a general practitioner reported lower use, continuing education programs might target those physicians.[16,17] In addition, public education programs should consider encouraging women themselves to initiate discussions of mammography with their physicians.

It is also not surprising that having had a clinical breast exam in the past 12 months was related to greater use. Apparently, a woman who undergoes a physical is more likely to be encouraged to have a mammogram. However, there continues to be considerable "missed referral opportunity"[4]: 28% of the women who had had a clinical breast exam in the previous year did not go on to get a mammogram (Table 1). In addition, many women reported not having had a clinical breast exam in the previous year (Table 1). Attention to both mammography and clinical breast exams is therefore important. Providers should continue to initiate discussion of screening and public education programs should continue to emphasize the importance of periodic checkups that include clinical breast exams and mammogram referrals, particularly for women of lower educational levels. □

Acknowledgments

This study was supported by the Division of Cancer Control, National Cancer Institute, ROI-CA44990-04.

We thank Marianne Raimondo and Sharon Dorfman for their contributions and Harriet Cramer for manuscript preparation.

References

1. Cancer statistics. *CA.* 1989;39:3.
2. Vanchieri, C. Medical groups' message to women: if 40 or older get regular mammograms. *J Natl Cancer Inst.* 1989;81:1126–1128.
3. *Healthy People 2000.* Washington, DC: US Dept of Health and Human Services; 1990.
4. NCI Breast Cancer Screening Consortium. Screening mammography—a missed clinical opportunity? *JAMA.* 1990;264:54–58.
5. Aday LA, Anderson RA. A framework for the study of access to medical care. *Health Serv Res.* 1974;9:208–220.
6. Lairson DR, Swint JM. Estimates of preventive versus nonpreventive medical care demand in an HMO. *Health Serv Res.* 1979;14:33–43.
7. Green LW, Kreuter MW. *Health Promotion Planning: An Educational and Environmental Approach.* Mountain View, Calif: Mayfield Publishing Co; 1991.
8. Hosmer DW, Lemeshow S. *Applied Logistic Regression.* New York, NY: John Wiley & Sons, Inc; 1989.
9. Centers for Disease Control. Trends in screening mammograms for women 50 years of age and older—behavioral risk factors surveillance system, 1987. *MMWR.* 1989;38:157–160.
10. Zapka J, Stoddard A, Costanza M, Greene H. Breast cancer screening by mammography: utilization and associated factors. *Am J Public Health.* 1989;79:1499–1502.
11. Zapka J, Stoddard A, Barth R, Costanza ME, Mas E. Breast cancer screening utilization by Latina community health center clients. *Health Educ Res.* 1989;4:461–468.
12. Dutton D: Financial organizational and professional factors affecting health care utilization. *Soc Sci Med.* 1989;23:721–735.
13. Lerman C, Rimer B, Trock B, Balshem A, Engstrom PF. Factors associated with repeat adherence to breast cancer screening. *Prev Med.* 1990;19:279–290.
14. Fox SA, Murata PJ, Stein JA. The impact of physician compliance on screening mammography for older women. *Arch Intern Med.* 1991;151:50–56.
15. Burg MA, Lane DS, Polednak AP. Age group differences in the use of breast cancer screening tests: the effects of health care utilization and socioeconomic variables. *J Aging Health.* 1990;2:514–530.
16. Albanes D, Weinberg GB, Boss L, Taylor PR. A survey of physicians' breast cancer early detection practices. *Prev Med.* 1988; 17:643–652.
17. American Cancer Society. 1989 survey of physicians' attitudes and practices in early cancer detection. *CA.* 1990;40:77–101.

Current Cigarette Smoking and Risk of Acute Pelvic Inflammatory Disease

Delia Scholes, PhD, Janet R. Daling, PhD, and Andy S. Stergachis, PhD

ABSTRACT

Objectives. Further information is needed on modifiable factors associated with the occurrence of acute pelvic inflammatory disease (PID). Cigarette smoking has been implicated as a risk factor for PID sequelae, but the association between smoking and PID has yet to be fully examined.

Methods. We conducted a population-based case-control study to evaluate smoking as a risk factor for acute PID. The case patients (n = 131) were women health maintenance organization (HMO) enrollees between the ages of 18 and 40 years who were treated for a first episode of PID. The control patients (n = 294) were randomly selected from the HMO enrollment files.

Results. Relative to never smokers, current smokers were at increased risk of PID. Women who smoked 10 or more cigarettes per day had a higher risk than did those who smoked less. Available data indicate that smoking status is not serving as a marker for uncontrolled confounding by lifestyle factors.

Conclusions. Our study results suggest that smoking represents a modifiable risk factor for acute PID. (*Am J Public Health.* 1992;82:1352–1355)

Introduction

Cigarette smoking has been implicated as a risk factor for ectopic pregnancy[1–5] and tubal infertility.[6] One way in which smoking may affect both of these outcomes is by increasing a woman's risk for pelvic inflammatory disease (PID), a severe upper genital tract infection that may result in scarring and obstruction of the fallopian tubes. Although one study[7] found an increased risk for smokers, the relationship between smoking and PID has yet to be fully examined. We evaluated this association in a population-based series of cases and controls from the Group Health Cooperative of Puget Sound.

Methods

The case patients consisted of women aged 18 to 40 years who experienced their first episode of acute PID between September 1984 and November 1985. The Group Health Cooperative automated hospital information system was used to select inpatients who had received a primary discharge diagnosis of acute PID or salpingitis. The automated pharmacy database was used to select potential ambulatory care PID cases. Medical records of women who were prescribed either doxycycline or a high-dose regimen of tetracycline[8] were reviewed for a diagnosis of PID. All medical records were reviewed for the presence of three clinical inclusion criteria: abdominal pain of 1 month's duration or less, cervical motion or uterine tenderness, and adnexal tenderness. Patients were included only if all three symptoms and signs were noted.

The control patients were part of a control group for a concurrent population-based study of ectopic pregnancy conducted at Group Health Cooperative. These women were selected randomly from the enrollment files. Control patients with reference dates (dates assigned to approximate the date of conception of the assigned case) between January 1984 and August 1986 were selected randomly from the enrollment files and matched to the women with ectopic pregnancies on age and county of residence.

The following inclusion criteria were applied to all potential research subjects: resident of King County; no prior history of PID; sexually active within the preceding year; no prior hysterectomy or bilateral salpingo-oophorectomy; not pregnant; no pelvic surgery or uterine instrumentation within the preceding 6 weeks; no vasectomized partner at reference date.

An in-person interview was used to collect information on cigarette smoking status at diagnosis/reference date, duration and amount smoked, and, for former smokers, length of time since quitting. Similar information was obtained on respondents' partners where applicable.

Unconditional logistic regression[9] was used to assess the role of smoking as

Delia Scholes is with the Center for Health Studies, Group Health Cooperative of Puget Sound, Seattle, Wash. Janet R. Daling and Andy S. Stergachis are with the Department of Epidemiology, School of Public Health and Community Medicine, University of Washington, Seattle. Andy S. Stergachis is also with the University of Washington's Department of Pharmacy and School of Pharmacy, where he is a Burroughs Wellcome Scholar.

Requests for reprints should be sent to Delia Scholes, PhD, Center for Health Studies, Group Health Cooperative of Puget Sound, 1730 Minor Ave, Suite 1600, Seattle, WA 98101.

This paper was submitted to the Journal May 10, 1991, and accepted with revisions May 13, 1992.

a risk factor while controlling for the confounding effects of age, race, gravidity, and number of sexual partners in the year prior to reference date. Risk estimates of smoking in relation to PID were not substantially altered when income, education, marital status, presence and type of contraception at reference date, and recent douching were added to the model.

Results

We interviewed 131 (72.4%) of the 181 case patients identified as eligible. Twelve percent of the 131 were hospitalized. Of the 428 women identified as potential controls, 314 (73.4%) were interviewed. We excluded 20 control patients with a history of PID.

Case patients were more likely to be younger, to have lower incomes and less education, to be non-White, to have been previously pregnant, and to have had more than one sex partner in the previous year (Table 1). They were less likely to be married at reference date.

More case patients (43.5%) than control patients (25.6%) were current smokers (Table 2). Current smokers experienced a twofold increase in risk (adjusted odds ratio [OR] = 2.1, 95% confidence interval (CI) = 1.3, 3.6) compared with never smokers. Women who smoked an average of 10 or more cigarettes per day were at higher risk than those who smoked less (OR = 2.4, 95% CI = 1.4, 4.2 vs OR = 1.2, 95% CI = 0.5, 3.0). The risks associated with current smoking were consistently higher among women with two or more recent sexual partners (Table 3). The percentage of current smokers did not differ between partners of case and partners of control patients (35% and 30%, respectively).

TABLE 1—Demographic Characteristics of Pelvic Inflammatory Disease Case and Control Patients

	Total Sample				Currently Smoking			
	Case Patients (n = 131)		Control Patients (n = 294)		Case Patients (n = 57)		Control Patients (n = 76)	
	No.	%	No.	%	No.	%	No.	%
Age, y								
18–24	60	45.8	76	25.9	28	46.7	18	23.7
25–29	31	23.7	75	25.5	12	38.7	12	16.0
30–34	22	16.8	86	29.3	10	45.5	24	27.9
35+	18	13.7	57	19.4	7	38.9	22	38.6
Race								
White	101	77.1	253	86.1	49	48.5	66	26.1
Black	23	17.6	23	7.1	8	34.8	7	30.4
Other	7	5.4	18	6.1	0	...	3	16.7
Income[a]								
<$15 000	55	42.3	57	30.6	26	47.3	16	28.1
$15 000–30 000	47	36.2	122	37.9	22	46.8	30	24.6
>$30 000	28	21.5	114	31.5	8	28.6	29	25.4
Education								
≤High school	58	44.3	69	28.0	33	56.9	28	40.6
>High school	73	55.7	225	72.0	24	32.9	48	21.3
Current marital status								
Never married	10	7.6	32	15.6	4	40.0	10	31.3
Married	34	26.0	164	44.5	15	44.1	39	23.8
Separated/divorced	4	3.1	15	3.8	3	75.0	5	33.3
Living as married	39	22.1	27	10.0	12	41.4	9	33.3
Regular partner	54	41.2	56	26.0	23	42.6	13	23.2
Gravidity								
Never	37	28.2	111	50.2	13	35.1	25	22.5
Ever	94	71.8	183	49.8	44	46.8	51	27.9
Recent douching[b]								
No	74	57.7	218	77.1	31	41.3	52	23.9
Yes	55	42.3	74	22.9	26	47.3	24	32.4
Current contraceptive method[c]								
None	24	18.5	58	18.6	11	45.8	16	27.6
Oral contraceptive	45	34.6	87	37.4	18	40.0	23	26.4
IUD	8	6.2	13	3.5	5	62.5	4	30.8
Barrier	29	22.3	86	28.3	11	37.9	16	18.6
Tubal ligation	15	11.5	38	8.9	8	53.3	15	39.5
Other	9	6.9	12	3.3	3	33.3	2	16.7
Number of sex partners in preceding 12 months[d]								
One	87	66.4	237	80.6	31	35.6	55	23.2
Two	22	16.8	32	13.7	10	45.5	14	43.8
Three or more	22	16.8	22	9.7	16	72.7	7	31.8

Note. Control percentages were adjusted to the age distribution of cases.
[a]One case and one control patient refused to provide this information.
[b]Data missing for two case and two control patients.
[c]Data missing for one case patient.
[d]Data missing for three control patients.

Discussion

Although we were able to represent a broader spectrum of disease by including outpatients and were able to use a population-based series of control patients, our data have certain limitations. Diagnostic misclassification was unavoidable because laparoscopic confirmation was not done in the majority of cases. Also, we could not obtain complete microbiological data, which prevented us from examining how types of infection might modify the association with smoking.

Since many of the variables associated with increased risk of PID are also associated with smoking,[10,11] the possibility remains that smoking is a marker for lifestyle attributes inadequately adjusted for in the analysis. Several factors argue against this possibility. First, although lifestyle may be crudely assessed in the available variables, the proportion of women who were current smokers was higher for cases than for controls in virtually all categories of each variable (Table 1). Second, the risk of PID was elevated only among current smokers, which is consistent with expectations for an acute infectious process; among current smokers, moreover, risk increased with increasing amount smoked and with increasing potential for exposure to a variety of infecting organisms (more than one recent sex partner).

TABLE 2—Current Cigarette Smoking and Risk of Pelvic Inflammatory Disease

	Case Patients (n = 131)		Control Patients (n = 294)			
	No.	%	No.	%	Odds Ratio[a]	95% CI
Never smoked	54	41.2	171	58.2	1.0	referent
Ever smoked	77	58.8	123	41.8	1.8	1.1, 2.8
Current smokers	57	43.5	76	25.9	2.1	1.3, 3.6
<10 per day	11	8.4	21	7.1	1.2	0.5, 3.0
≥10 per day	46	35.1	55	18.7	2.4	1.4, 4.2
Former smokers	20	15.3	47	16.0	1.3	0.7, 2.4
Until 3 years ago	12	9.2	16	5.4	1.8	0.8, 4.3
>3 years ago	8	6.1	31	10.5	0.9	0.4, 2.2

[a]Adjusted for age, race, gravidity, and number of sex partners in the preceding 12 months.

TABLE 3—Current Cigarette Smoking and Risk of Pelvic Inflammatory Disease According to Number of Sex Partners in the Preceding Year

	One Partner					Two or More Partners				
	Case Patients (n = 72)		Control Patients (n = 200)		Odds Ratio[a]	Case Patients (n = 39)		Control Patients (n = 45)		Odds Ratio[a]
	No.	%	No.	%	(95% CI)	No.	%	No.	%	(95% CI)
Never smoked	41	56.9	145	72.5	1.0 (referent)	13	33.3	24	53.3	1.0 (referent)
Current smokers	31	43.1	55	27.5	1.9 (1.0, 3.4)	26	66.7	21	46.7	2.9 (1.1, 7.5)
<10 per day	3	4.2	12	6.0	1.1 (0.3, 4.4)	8	20.5	9	20.0	1.6 (0.4, 5.5)
≥10 per day	28	38.9	43	21.5	2.1 (1.1, 3.9)	18	46.2	12	26.7	3.9 (1.3, 11.2)

[a]Adjusted for age, race, gravidity.

Finally, we noted that the current smoking patterns of respondents' partners were unrelated to respondents' risk.

There are a number of plausible ways in which smoking may increase the risk of PID. The effect of nicotine on tubal ciliary function and ovum transport has been documented,[12] and the same impaired ciliary wave action that can fail to move the ovum toward implantation in the uterus may result in an inability to effectively repel ascending infectious organisms.

Studies have also provided evidence that smoking impairs the immune response.[13–19] A recent examination of the effect of cigarette smoking specifically on local cervical immunity found that current smokers had a decrease in the concentration of Langerhans' cells in the cervical epithelium.[20] Because these cells help detect and present antigens to circulating T-lymphocytes, a decrease could affect susceptibility to PID pathogens.

The only other study to assess this relationship used hospital-based data from the Women's Health Study[7]; the OR for current smokers compared to never smokers was 1.7. Our findings agree with findings from studies of PID sequelae, as well.[1–6] An association between smoking and urethritis in men, which is caused by many of the same organisms that cause PID,[21] has also been reported.[22]

More information is needed about modifiable noninfectious factors that may be associated with the acquisition or spread of cervical infection. This population-based study describes an association between one such factor—current smoking—and acute PID. We provide evidence that uncontrolled confounding is not the source of these findings and show that there is reason to believe that smoking itself may be acting to increase the risk of acute PID. □

Acknowledgments

This research was supported in part by grants 1-R01-HD17657 and 3-R01-HD17657-S1 from the National Institute of Child Health and Human Development, National Institutes of Health.

These findings were reported in part at the Society for Epidemiological Research 23rd Annual Meeting in Snowbird, Utah, June 1990.

The authors would like to acknowledge Noel Weiss, Kathleen Saunders, Barbara Young, Sandra Howard, Carol Ure, Jude Ballard, and the interviewing team that collected these data for their assistance with this project.

References

1. Campbell OM, Gray RH. Smoking and ectopic pregnancy: a multinational case-control study. In: Rosenberg M, ed. *Smoking and Reproductive Health.* Littleton, Mass: PSG Publishing Co; 1987.
2. Chow W-H, Daling JR, Weiss NS, Voigt L. Maternal cigarette smoking and tubal pregnancy. *Obstet Gynecol.* 1988;71:167–170.
3. Handler A, Davis F, Ferre C, Yeko T. The relationship of smoking and ectopic pregnancy. *Am J Public Health.* 1989;79:1239–1242.
4. Coste J, Job-Spira N, Fernandez H. Increased risk of ectopic pregnancy with maternal cigarette smoking. *Am J Public Health.* 1991;81:199–201.
5. Stergachis AS, Scholes D, Daling JR, Weiss NS, Chu J. Maternal cigarette smoking and the risk of ectopic pregnancy. *Am J Epidemiol.* 1991;133:332–337.
6. Daling JR, Weiss NS, Spadone L, Moore DE, Voigt L. Cigarette smoking and primary tubal infertility. In: Rosenberg M, ed. *Smoking and Reproductive Health.* Littleton, Mass: PSG Publishing Co; 1987.
7. Marchbanks PA, Lee NC, Peterson HB. Cigarette smoking as a risk factor for pelvic inflammatory disease. *Am J Obstet Gynecol.* 1990;162:639–644.
8. US Department of Health and Human Services. Sexually transmitted diseases: treatment guidelines, 1982. *MMWR.* 1982; 31(Suppl):43S.
9. Breslow NE, Day NE. *Statistical Methods in Cancer Research.* Vol 1. Lyon, France: IARC; 1980.
10. Novotny TE, Warner KE, Kendrick JS, Remington PL. Smoking by blacks and whites: socioeconomic and demographic differences. *Am J Public Health.* 1988;78: 1187–1189.
11. O'Malley PM, Bachman JG, Johnston LD. Period, age, and cohort effects on substance use among young Americans: a decade of change, 1976–86. *Am J Public Health.* 1988;78:1315–1189.
12. Weathersbee PS. Nicotine and its influence on the female reproductive system. *J Reprod Med.* 1980;25:243–250.
13. Thomas W, Holt PG, Keast D. Effect of cigarette smoking on primary and secondary humoral responses of mice. *Nature.* 1973;243:240–241.
14. Rozman TL, Rogers AS. The immunosuppressive potential of products derived from cigarette smoke. *Am Rev Resp Dis.* 1973; 108:1158–1163.
15. Schwartz SL. Interaction of nicotine and other amines with the endocytic and exo-

cytic functions of macrophages. *Fed Proc (Biochemical Aspects of Toxic Agents).* 1976;35:85–88.
16. Burton RC. Smoking, immunity, and cancer. *Med J Aust.* 1983;2:411–412.
17. Winkel P, Statland BE. The acute effect of cigarette smoking on the concentrations of blood leukocyte types in healthy young women. *Am J Clin Pathol.* 1981;75:781–785.
18. Chalmer J, Holt PG, Keast D. Cell-mediated immune responses to transplanted tumors in mice chronically exposed to cigarette smoke. *J Natl Cancer Inst.* 1975;55:1129–1134.
19. Andersen P, Pedersen OF, Bach B, Bonde GJ. Serum antibodies and immunoglobulin in smokers and nonsmokers. *Clin Exp Immunol.* 1982;47:467–473.
20. Barton SE, Jenkins D, Cuzick J, Maddox PH, Edwards R, Singer A. Effect of cigarette smoking on cervical epithelial immunity: a mechanism for neoplastic change? *Lancet.* 1988;ii(8612):652–654.
21. Thompson SE, Washington AE. Epidemiology of sexually transmitted chlamydia trachomatis infections. *Epidemiol Rev.* 1983;5:96–122.
22. Pessione F, Dolivo M, Casin I, et al. Sexual behavior and smoking: risk factor for urethritis in men. *Sex Trans Dis.* 1988;15:119–122.

Dr. Derrick B. Jelliffe to Receive Posthumous Award at APHA Annual Meeting

Dr. Derrick B. Jelliffe will posthumously receive the "Lifetime Achievement Award" from APHA's International Health Section at this year's APHA meeting in Washington, DC, November 8 to 12. (Dr. Jelliffe, a renowned pediatrician and advocate for international maternal and child health and nutrition, as well as professor emeritus at the University of California–Los Angeles School of Public Health, died March 18 at the age of 71.) The tribute will take place Tuesday, November 10, from 7:00 to 8:30 PM. Dr. Susi Kessler, chair of the International Health Section, will present the award. Dr. Michael C. Latham, professor and director of the Program in International Nutrition, Cornell University, will deliver the tribute's keynote address, "Derrick Jelliffe: Pioneer in Child Health and Infant Feeding." A reception will follow.

Lead Exposure and Child Behavior

ABSTRACT

Objectives. Unlike cognitive impairments associated with lead exposure, lead-associated child behavior problems have been difficult to specify, particularly in young children.

Methods. The Child Behavior Checklist (CBCL) and the Center for Epidemiologic Studies Depression Scale were used as the outcome and confounding variables, respectively, of major interest. These measures were examined with respect to blood lead levels of 201 African-American children aged 2 through 5 years.

Results. In comparison with the low exposed group, the high exposed group (two consecutive blood lead levels ≥15 μg/dL) had a significantly higher mean CBCL Total Behavior Problem Score (TBPS) and Internalizing and Externalizing scores; when other factors, including maternal depressive symptomatology, were controlled for, regression procedures indicated a .18-point TBPS increase for each unit increase in lead and a 5.1-point higher TBPS in the high exposed group; children in this group were 2.7 times more likely to have a TBPS in the clinical range.

Conclusions. Through its use of a standardized parent-report measure of behavior and its consideration of maternal morale in multiple linear and logistic regression procedures, this study provides further evidence of lead's detrimental effect on child behavior at levels typical of present-day exposure. (*Am J Public Health.* 1992;82:1356–1360)

William G. Sciarillo, ScD, MSN, Greg Alexander, ScD, MPH, and Katherine P. Farrell, MD, MPH

Introduction

Lead poisoning is the most important pediatric environmental health problem in the United States.[1] Disturbances in neurological and systemic functions have been identified at levels that were once thought not to be cause for concern (i.e., 10 to 25 μg/dL).[2] There is strong evidence that low-level exposure impairs cognitive development,[3] and long-term effects of childhood exposure have been reported.[4]

Lead-associated behavioral difficulties in young children are insidious and difficult to specify. Nonspecific symptoms (e.g., overactivity, irritability, lethargy) may not be recognized until the child enters school. In this study we examine the influence of lead exposure on maternal-reported behavior in children aged 2 through 5 years.

Because maternal morale may color salient aspects of the early caregiving environment, affecting lead exposure (related to compromised caretaking),[5–8] behavioral development,[9] and maternal perceptions of child behavior,[10] maternal depressive symptomatology is considered in this study as a potentially important confounding factor. Low-income women who live alone with their children are at greater risk for depression.[11,12]

Methods

A convenience sample of 150 mother-child dyads from the Baltimore Soil Lead Abatement Demonstration Project's clinic were enrolled in this study during the period from February through April 1989; 54 dyads from the Kennedy Krieger Institute's Lead Poisoning Referral Center were recruited at their clinic visits in April, May, and June 1989. Recruitment from both a population-based study and a clinical facility for referred children ensured a broad range of lead exposure. All subjects met the following enrollment criteria: (1) the child was between 2 and 5 years old; (2) the mother or female guardian accompanied the child to the clinic; and (3) both were African-American. There were three refusals to participate (Kennedy Institute), resulting in a sample of 201 dyads.

Lead concentrations were determined from venous samples collected during the clinics. Children with a blood lead level greater than or equal to 15 μg/dL on two consecutive occasions—at the clinic appointment immediately prior to enrollment in this study and at the first appointment during the study period—were considered the high exposed group. Consistent with preestablished criteria,[13] children with a free erythrocyte protoporphyrin level greater than or equal to 35 μg/dL and a blood lead level less than 25 μg/dL were considered iron-deficient (in the absence of lead poisoning).

Child behavior was measured by interviewing the mother during the child's clinic visit. The Child Behavior Checklist

At the time of the study, William G. Sciarillo was with Children's Medical Services, Local and Family Health Administration, Department of Health and Mental Hygiene, Baltimore, Md. Greg Alexander is with the Maternal and Child Health Program, School of Public Health, University of Minnesota, Minneapolis, Minn. Katherine P. Farrell is with Community Health Services, Anne Arundel County Health Department, Annapolis, Md.

Requests for reprints should be sent to William G. Sciarillo, ScD, Executive Director, Association for the Care of Children's Health, 7910 Woodmont Ave, Suite 300, Bethesda, MD 20814.

This paper was submitted to the Journal June 18, 1991, and accepted with revisions April 27, 1992.

(CBCL),[14–16] a standardized scale designed to quantify troublesome parent-reported behaviors, was used. The CBCL is widely accepted as a reliable and valid approach to behavioral assessment.[17]

Maternal depressive symptomatology was measured by interview; the Center for Epidemiologic Studies Depression Scale (CES–D),[18] a 20-item index of the number and frequency of depressive symptoms,[19] was used. A CES-D score of 16 or more is clinically significant (significance has been validated by correlation with clinical ratings of depression and by associations with negative life events) and may also indicate nonspecific psychological distress.[20,21]

Results

There were 123 children in the low exposed (LE) group and 78 in the high exposed (HE) group. Mean blood lead levels prior to this study were 11.3 μg/dL (SD = 4.3) for the LE group and 28.6 μg/dL (SD = 9.3) for the HE group; mean levels during the study were 9.2 μg/dL (SD = 2.9) and 27.8 μg/dL (SD = 10.4), respectively.

Mean CBCL Total Behavior Problem Scores (TBPSs), Internalizing and Externalizing T-scores, based upon the percentile ranking of raw scores, are presented in Table 1. The HE group consistently had more maternal-reported troublesome behaviors. Prevalence of TBPSs, Internalizing, and Externalizing T-scores in the clinical range (90th percentile) are presented in Table 2. The proportion of HE children with clinical T-scores was approximately double that of the LE group.

The prevalence of subscale T-scores in the clinical range (98th percentile) was examined, providing a more specific perspective of maternal-reported behavior problems. Overall, 17.9% of LE children and 30.8% of HE children had one or more subscales in the clinical range (odds ratio = 2.0; 95% CI = 1.0, 4.0).

The prevalence of clinical T-scores for each subscale of the 2- to 3-year-old (both sexes) profile (n = 118), the 4- to 5-year-olds boys' profile (n = 46), and the 4- to 5-year-old girls' profile (n = 37) is presented in Figures 1 and 2. This prevalence was generally higher in the HE group across all three profiles. For 2- to 3-year-olds and for 4- to 5-year-old boys, this tendency was particularly evident toward the Externalizing end of the continuum of the profiles (e.g., for the Aggressive subscale).

TABLE 1—Mean T-Scores of Low Exposed and High Exposed Groups on the Child Behavior Checklist (CBCL)

	Low Exposed (n = 123)		High Exposed (n = 78)	
CBCL Score	Mean	(SD)	Mean	(SD)
Total*	51.0	(10.5)	55.8	(12.5)
Internalizing**	51.0	(10.2)	55.2	(10.2)
Externalizing***	51.0	(10.0)	54.8	(12.0)

*$t = -2.93$; $df = 199$; $P < .01$.
**$t = -2.89$; $df = 199$; $P < .01$.
***$t = -2.46$; $df = 199$; $P = .01$.

TABLE 2—Prevalence of Clinical T-Scores for Low Exposed and High Exposed Groups on the Child Behavior Checklist (CBCL)

	Low Exposed (n = 123)		High Exposed (n = 78)	
CBCL Score	%	(No.)	%	(No.)
Total*	16.3	(20)	32.0	(25)
Internalizing**	10.6	(13)	19.2	(15)
Externalizing***	14.6	(18)	25.6	(20)

*$\chi^2 = 6.85$; $df = 1$; $P < .01$.
**$\chi^2 = 2.99$; $df = 1$; $P = .08$.
***$\chi^2 = 3.77$; $df = 1$; $P = .05$.

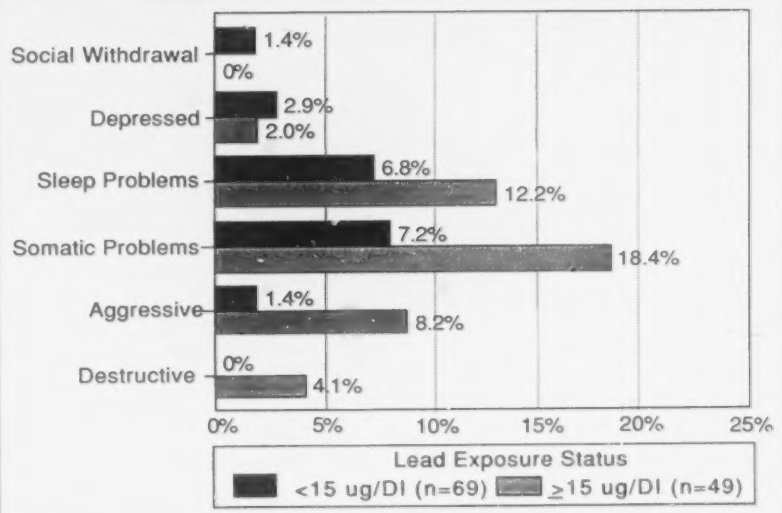


FIGURE 1—Prevalence of Child Behavior Checklist subscale T-scores in clinical range (98th percentile), 2- to 3-year-old boys and girls (n = 118).

Other factors that may have accounted for maternal-reported behavior differences between the two groups were examined. Mean age was 46.4 months (SD = 13.7) for LE children and 44.7 months (SD = 12.9) for HE children ($P = .38$). Maternal education averaged 12.2 years (SD = 1.8) for the LE group and 11.0 years (SD = 1.6) for the HE group ($P < .001$). Mean maternal CES-D scores were 11.3 (SD = 9.9) for the LE group and 13.2 (SD = 9.9) for the HE group ($P = .18$); prevalence of clinical CES-D scores was 25.2% and 34.6%, respectively ($P = .15$). HE children were more likely to come from households in which the mother was not a high school graduate ($P < .001$), was unemployed ($P < .01$), was single ($P < .05$), or was head of the household ($P < .01$), and from

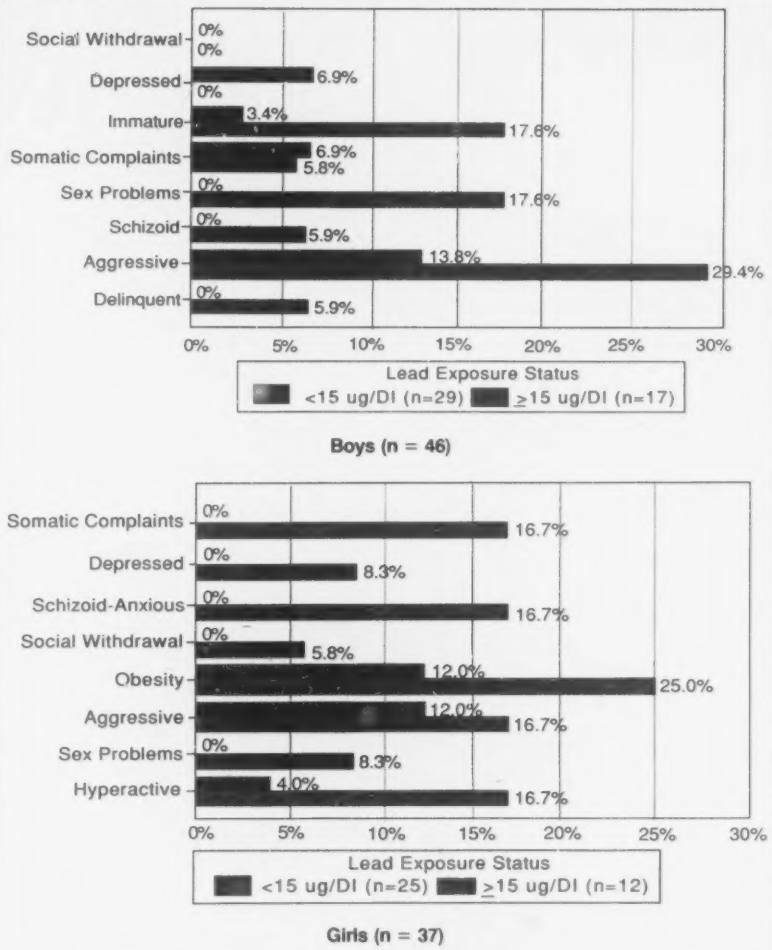


FIGURE 2—Prevalence of Child Behavior Checklist subscale T-scores in clinical range (98th percentile), 4- to 5-year-old boys and girls.

TABLE 3—Multiple Linear Regression on Child Behavior Checklist Total Behavior Problem Score

	Regression Coefficient	Standard Error	*P* Value
Maternal education, y	1.00	0.450	.0279
Maternal occupation: unemployed	2.15	1.614	.1841
Maternal marital status: not married	0.62	1.943	.7505
Preschool children in household: ≥4	−2.78	2.458	.2591
Maternal CES-D score	0.55	0.076	.0001
Child's age, m	0.14	0.054	.0128
Child's sex: male	0.71	1.483	.6334
Iron deficiency[a]	2.95	2.455	.2308
Blood lead, μg/dL	0.18	0.073	.0125

Note. $F = 7.82$, $r^2 = 0.269$.
Intercept = 22.97.

[a]Iron deficiency in the absence of lead poisoning (i.e., free erythrocyte protoporphyrin ≥35 μg/dL *and* blood lead <25 μg/dL).

households in which there was an increased likelihood of four or more preschool-aged children in the household ($P < .05$).

When all other variables in the model were simultaneously controlled for through multiple linear regression, a significant increase ($P = .01$) in the TBPS with increasing blood lead levels was found (Table 3). The magnitude of this effect was quite small (the TBPS increased by .18 points for each unit increase in lead). However, when a dichotomous measure of blood lead (LE vs HE group) was substituted in the model and all other variables in the model were controlled for, a 5.1-point higher TBPS was found in HE children ($P = .001$). Maternal CES-D score was the most significant influence on the child's TBPS ($P = .0001$); child's age and maternal education also had significant independent influence. Similar results were found for the CBCL Internalizing and Externalizing scores.

Multiple logistic regression indicated that HE children were 2.7 times more likely to have a TBPS in the clinical range (Table 4). Maternal depressive symptomatology was most influential (odds ratio = 4.7).

Discussion

Young children with blood lead levels greater than or equal to 15 μg/dL on two consecutive occasions presented more maternal-reported maladaptive behaviors than did children with lower blood lead levels. Initial comparisons between the LE and HE groups indicated significantly higher mean TBPSs, Internalizing scores, and Externalizing scores in HE children. These children also had a prevalence of such scores in the clinical range (90th percentile) that was nearly two times higher than that of children in the LE group.

The HE group generally had a higher prevalence of CBCL subscale scores in the clinical range (98th percentile), especially at the Externalizing end of the continuum of the profiles for 2- to 3-year-old children of both sexes and for 4- to 5-year-old boys. Four- to five-year-old girls in the HE group had a higher prevalence of clinical subscales over the entire Internalizing-Externalizing continuum. The subscales represent factor analytically derived behavioral descriptions, not diagnostic inferences. The prevalence of clinical range subscales for LE children was quite similar to those found in previous reports for young children without known biological risk factors.[15(pp 64,65),22]

TABLE 4—Multiple Logistic Regression on Child Behavior Checklist Total Behavior Problem Score in Clinical Range (90th Percentile)

	Model Level	Adjusted Odds Ratio	95% Confidence Interval
Maternal education, y	<12	0.7	0.3, 1.6
Maternal occupation	Unemployed	1.2	0.5, 2.6
Maternal marital status	Not married	2.0	0.6, 6.4
Preschool children in household	≥4	0.2	0.0, 1.0
Maternal CES-D score	≥16	4.7	2.2, 10.1
Child's age, y	4 or 5	0.9	0.4, 1.8
Child's sex	Male	1.1	0.5, 2.4
Iron deficiency only	Yes	0.8	0.2, 3.4
Blood lead group	High exposed	2.7	1.2, 5.7

Note. Likelihood ratio $\chi^2 = 32.10$; $df = 9$; $P < .001$.

The effect of lead on externalizing behaviors (e.g., increased motor activity) has been demonstrated in animal models.[23–25] The association between exposure and "hyperactive" child behavior has been reported in several case-control studies[26–28] and in Needleman et al.'s population-based study, in which a consistent dose-response relationship between lead and teacher-reported nonadaptive behaviors was found.[29] Others reported children with increased lead levels as demonstrating more behavior problems, particularly with respect to conduct, inattentiveness, passivity, and hyperactivity.[30] A related finding is that hyperactive children have been found to have higher urinary lead levels after penicillamine challenge.[31]

When the influence of selected factors was controlled for in this study through multiple regression analyses, the magnitude of the lead-behavior effect was such that higher lead exposure was associated with a 5.1-point increase in the TBPS. Although this association may not appear to be clinically relevant for an individual child, an overall 5-point upward shift in the normal distribution of TBPSs would have resulted in almost 82% more children scoring in the clinical range. Also, the HE group was 2.7 times more likely to have a TBPS in the clinical range. This finding takes on particular significance in view of the pandemic nature of lead exposure at levels greater than or equal to 15 μg/dL in young children.[2]

The most influential factor on the TBPS in both the multiple linear and logistic regression procedures was the maternal CES-D score. This influence can be explained in at least five ways: (1) mother and child share a similar genotype, (2) mother and child may be reacting to a common factor (e.g., poverty), (3) increased maternal depressive symptoms detrimentally influence child behavior, (4) child behavior problems detrimentally influence maternal morale, and (5) mothers with lower morale rate their children's behavior more negatively. It is reasonable to maintain that some indistinct dynamic combination of all five is operational here. Although the cross-sectional design precludes a more specific delineation, including the maternal depressive symptomatology variable in the regression models provided some degree of empirical control for these possible influences.

With respect to the later possible influence, Conrad and Hammen investigated the relationship of maternal depression to perceptions of internalizing and externalizing disorders in children. They reported that maternal depressive symptomatology and psychiatric clinical status were not associated with misperceptions of child behavior problems.[32] Using teacher ratings as an independent criterion for rating the accuracy of maternal ratings, Richters and Pellegrini found no evidence that depressed mood, depressed clinical state, or history of clinical depression attenuated mother-teacher agreement about child behavior problems.[33]

In addition to the epidemiological significance of an increased prevalence of maternal-reported troublesome behaviors at lead levels typical of present-day exposure, a major clinical implication of this study is that young children with reported behavioral difficulties should undergo appropriate lead screening. Also, the care context in which lead evaluation and treatment occur provides an opportunity to address parental concerns about child behavior. Appropriate interventions, including changing parental misconceptions about normal child behavior, can be initiated. The parent-child dyad is often a captive audience during the lead evaluation and treatment process, and the parent may be more receptive to interventions at this time.

The results of this study, with its particular attention to maternal-reported maladaptive behaviors in young children, lends support to a belief that undue lead exposure in early childhood may have a pervasive influence on the prevalence of juvenile delinquency in this country. □

Acknowledgments

Parts of this study were supported through the Baltimore Soil Lead Abatement Demonstration Project/Maryland Department of the Environment and through the Kennedy Kreiger Institute, Baltimore, Md.

The authors wish to thank Alan Ross, PhD, Vince Smeriglio, PhD, and Mark Farfel, ScD, for their assistance in the design, conduct and analysis of the study, and Julian Chisolm, MD, for his review and discussion of the manuscript (all affiliated with Johns Hopkins University, Baltimore, Md). We also thank Merrill Brophy, RN, MSN, and Marjorie Sheehan, RN, for their assistance in coordinating data collection at, respectively, the Baltimore Soil Lead Abatement Demonstration Project and the Kennedy Kreiger Institute sites.

References

1. Bellinger D, Leviton A, Waternaux C, Needleman H, Rabinowitz M. Low-level lead exposure, social class, and infant development. *Neurotoxicol Teratol.* 1989;10:497–503.
2. *The Nature and Extent of Lead Poisoning in Children in the United States: A Report to Congress.* Washington, DC: Agency for Toxic Substances and Disease Registry; 1988. DHHS publication 99-2966.
3. Needleman HL, Gatsonis CA. Low-level lead exposure and the IQ of children. *JAMA.* 1990;263:673–678.
4. Needleman HL, Schell A, Bellinger D, Leviton A, Allred E. The long-term effects of exposure to low doses of lead in childhood: an 11-year follow-up report. *N Engl J Med.* 1989;322:83–88.
5. Milar CR, Schroeder SR, Mushak P, Dolcourt JL, Grant LD. Contributions of the caregiving environment to increased lead burden of children. *Am J Ment Defic.* 1980;84:339–344.
6. Hunt TJ, Hepner R, Seaton KW. Childhood lead poisoning and inadequate child care. *Am J Dis Child.* 1982;136:538–542.
7. Dietrich KN, Krafft KM, Pearson DT, et al. Contribution of social and developmental factors to lead exposure during the first year of life. *Pediatrics.* 1985;75:1114–1119.
8. Panaccione VF, Wahler RG. Child behavior, maternal depression, and social coercion as factors in the quality of child care. *J Abnorm Child Psychol.* 1986;14:263–278.

9. Zuckerman BS, Beardslee WR. Maternal depression: a concern for pediatricians. *Pediatrics.* 1987;79:110–117.
10. Brody GH, Forehand R. Maternal perceptions of maladjustment as a function of the combined influence of child behavior and maternal depression. *J Consult Clin Psychol.* 1986;54:237–240.
11. Rosenfield S. Sex differences in depression: do women always have higher rates? *J Health Soc Behav.* 1980;21:33–42.
12. Orr ST, James S. Maternal depression in an urban pediatric practice: implications for health care delivery. *Am J Public Health.* 1984;74:363–365.
13. Yip R, Schwartz S, Deinard AS. Screening for iron deficiency with the erythrocyte protoporphyrin test. *Pediatrics.* 1983;72: 214–219.
14. Achenbach TM. The Child Behavior Profile: an empirically based system for assessing children's behavioral problems and competencies. *Int J Ment Health.* 1979;7: 24–42.
15. Achenbach TM, Edelbrock CS. *Manual for the Child Behavior Checklist and the Revised Behavior Profile.* Burlington, Vt: University Associates in Psychiatry: 1983.
16. Achenbach TM, Edelbrock CS, Howell CT. Empirically based assessment of the behavioral/emotional problems of 2- and 3-year-old children. *J Abnorm Child Psychol.* 1987;15:629–650.
17. Achenbach TM, Brown JS. *1989 Bibliography of Published Studies Using the Child Behavior Checklist and Related Materials.* Burlington, Vt: University of Vermont; 1989.
18. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1: 385–401.
19. Roberts RE. Reliability of the CES-D Scale in different ethnic contexts. *Psychol Res.* 1980;2:125–134.
20. Weissman MM, Sholomskas D, Pottenger M, Prusoff BA, Locke BZ. Assessing depressive symptoms in five psychiatric populations: a validation study. *Am J Epidemiol.* 1977;106:203–214.
21. Vernon SW, Roberts RE. Measuring nonspecific psychological distress and other dimensions of psychopathology. *Arch Gen Psychiatry.* 1981;38:1239–1247.
22. Larson CP, Piess IB, Miettinen O. Preschool behavior disorders: their prevalence in relation to determinants. *J Pediatr.* 1988; 113:278–285.
23. Sauerhoff M, Michaelson IA. Hyperactivity and brain catecholamines in lead-exposed developing rats. *Science.* 1971;173: 1022–1024.
24. Silbergeld EK, Hruska RE. Neurochemical investigations of low level lead exposure. In: Needleman HL, ed. *Low Level Lead Exposure.* New York, NY: Raven Press; 1980.
25. Ferguson SA, Bowman RE. Effects of postnatal lead exposure on open field behavior in monkeys. *Neurotoxicol Teratol.* 1989;12:91–97.
26. David O, Clark J, Voeller K. Lead and hyperactivity. *Lancet.* 1972;2:900–903.
27. David OJ, Hoffman SP, Sverd J, Clark J, Voeller K. Lead and hyperactivity. Behavioral response to chelation: a pilot study. *Am J Psychiatry.* 1976;133:1155–1158.
28. David O, Clark J, Hoffman S. Childhood lead poisoning: a re-evaluation. *Arch Environ Health.* 1979;34:106–111.
29. Needleman HL, Gunnoe C, Leviton A, et al. Deficits in psychologic and classroom performance of children with elevated dentine lead levels. *N Engl J Med.* 1979;300: 689–695.
30. Lansdown R, Yule W, Urbanowicz M, Miller I. Blood lead, intelligence, attainment and behavior in school children: overview of a pilot study. In: Rutter M, Jones RR, eds. *Lead versus Health.* New York, NY: John Wiley and Sons; 1983.
31. Gittelman R, Eskenazi B. Lead and hyperactivity revisited: an investigation of nondisadvantaged children. *Arch Gen Psychiatry.* 1983;40:827–832.
32. Conrad M, Hammen C. Role of maternal depression in perceptions of child maladjustment. *J Consult Clin Psych.* 1989;57: 663–667.
33. Richters J, Pellegrini D. Depressed mothers' judgements about their children: an examination of the depression-distortion hypothesis. *Child Dev.* 1989;60:1068–1075.

Pregnancy among the Hmong: Birthweight, Age, and Parity

Deborah Helsel, BA, RN, Diana B. Petitti, MD, and Peter Kunstadter, PhD

ABSTRACT

Objectives. **The influx of Southeast Asians into the United States allows for the study of this special population and contributes to a broader understanding of reproductive health.**

Methods. **We used information on birth certificates to identify 1937 Hmong children born 1985 through 1988 in California, and we compared birthweight and reproductive factors as related to these children with the same factors as related to 3776 White, non-Hispanic children born in the same period.**

Results. **Mean birthweight among Hmong children (3311 g) was significantly lower ($P < .05$) than among White children (3452 g), but the proportion of births under 1500 g was higher for Whites. Hmong women were of much higher parity and were more likely to deliver at both a young (< 18 years) and an old (> 40 years) maternal age. At every age and every parity, however, Hmong women had cesarean sections at one-half to one-tenth the rate of White women.**

Conclusions. **Despite a high proportion of births at high parity and advanced maternal age, Hmong women gave birth to very low-birthweight babies at essentially the same rates as White women. Their lower cesarean section rates, however, deserve further attention. (*Am J Public Health.* 1992;82:1361–1364)**

Introduction

The influx of large numbers of Southeast Asian refugees into the United States since 1975 has allowed for studies of the health status and practices of the various ethnolinguistic groups represented in the migration, which in turn have contributed to an understanding of the unique concerns of refugee populations and of reproductive health in general. The most recent wave of Southeast Asian immigration began in 1975 and included the Hmong, an agricultural people who live in relatively remote villages in the mountainous regions of southwestern China and northern Southeast Asia.[1]

As immigrants to the United States, the Hmong were more rural, less educated, less occupationally skilled, and less "Westernized" than earlier arrivals.[2] Most Hmong immigrants had little formal education and did not speak or read English, making communication in U.S. society difficult. Unemployment rates for newly arrived Hmong were well over 50%, and most (about 75%) of these immigrants depended on cash assistance. These factors indicate the difficulties the Hmong have had in adapting to life in the United States.[3]

The primary purpose of our study was to compare reproductive factors among Hmong immigrants with those among White, non-Hispanic women. Hmong women are unlikely to use tobacco or alcohol;[4] however, obstetrical care providers have encountered many other health care problems in Southeast Asian immigrant patients[4–7] that have previously been associated with poor pregnancy outcome. Such problems include low socioeconomic status and late prenatal care.[8,9] Although some of these risk factors among the Hmong might be expected to result in a high proportion of low-birthweight infants, this association has not been noted in related works.[5,10]

Methods

Merced and San Joaquin counties in the Central Valley of California were chosen for this study because of the large concentration of Hmong living there. Birth certificates filed 1985 through 1988 for Merced County and 1986 through 1988 for San Joaquin County were examined to identify singleton live births to Hmong and other Asian women. A random sample of singleton births to White, non-Hispanics in the same counties and years and equal in number to the Asian births was also selected. A total of 3776 Asian births and 3776 White, non-Hispanic births were identified.

To determine ethnicity within the Asian group, the names of both parents were inspected. Ethnicity was assigned using protocols designed with the help of members of the Hmong community. Distinctive spellings and syllable patterns, and lists of Southeast Asian clans and family naming systems, such as the one described in Muecke,[11] eventually made this a straightforward effort. Infants born of mixed marriages (of whom there were very few) were eliminated.

Deborah Helsel is with the Department of Social and Behavioral Sciences, Diana B. Petitti is with the Department of Family and Community Medicine, and Peter Kunstadter is with the Institute for Health Policy Studies at the University of California in San Francisco.

Requests for reprints should be sent to Dr. Diana B. Petitti, Department of Family and Community Medicine, AC-9, Box 0900, University of California-San Francisco, San Francisco, CA 94143.

This paper was submitted to the Journal December 30, 1991, and accepted with revisions March 12, 1992.

TABLE 1—Percentage Distribution of Birthweight in Singleton Births to Hmong and White Women: Merced and San Joaquin Counties, California

Ethnicity	Births, n	Birthweight Category <1000 g, %	1000–1499 g, %	1500–2499 g, %	2500+ g, %
Hmong	1937	0.1	0.2	4.3	95.4
White, non-Hispanic	3776	0.4	0.2	3.3	96.1

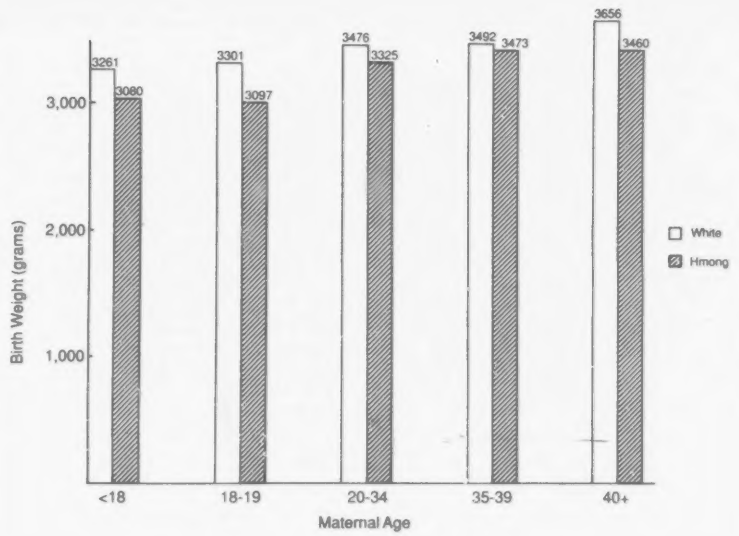


FIGURE 1—Mean birthweight in singleton births to Hmong and White women by maternal age.

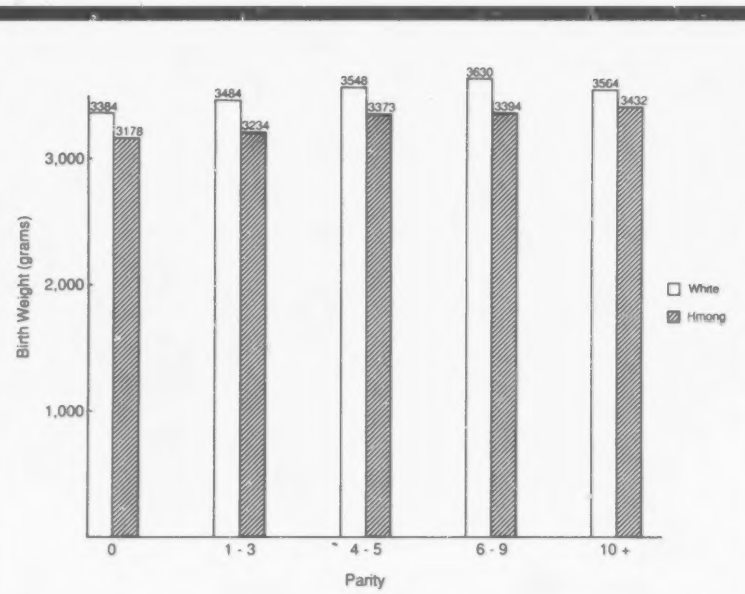


FIGURE 2—Mean birthweight in singleton births to Hmong and White women by parity.

The Asian births included 1937 Hmong. This group was compared with the 3776 White, non-Hispanic births. The statistical significance of differences in birthweight and gestational age were tested using the *t*-statistic for continuous variables. For polychotomous variables, the chi-square statistic was used.

Results

The mean birthweight in singleton Hmong births was 3311 g; in Whites, it was 3452 g ($P < .05$). The overall low-birthweight (< 2500 g) rate for the Hmong was 4.6%, not significantly ($P > .05$) different from that for the Whites (3.9%). Hmong women had a lower proportion of births weighing less than 1500 g than did White women (Table 1), but again, the difference was not significant ($P > .05$). The mean gestational age in the Hmong was 268.3 days; in Whites, it was 272.3 days ($P < .05$).

Parity and maternal age of Hmong mothers differed markedly from that of White mothers. As to parity, 36.6% of Whites and 11.3% of Hmong were primiparous; on the other hand, 41.3% of Hmong and only 2.4% of Whites were para 5 or greater ($P < .05$). As to maternal age, 5.0% of White mothers and 7.9% of Hmong mothers were under age 18 whereas 4.8% of Hmong mothers and 0.8% of White mothers were over age 40 ($P < .05$). Mean birthweight among the Hmong was lower at all parities and ages (Figures 1 and 2).

The overall rate of cesarean sections for Hmong women was 3.9% compared with 23.8% for White women ($P < .05$). The low Hmong rate was not due to age or parity (Table 2).

As a group, White mothers tended to begin prenatal care earlier than Hmong mothers; 71.9% of White mothers had begun care by the third month of pregnancy compared with 34.4% of Hmong mothers ($P < .05$). However, only 0.7% of the Hmong mothers received no prenatal care compared with 2.1% of the White mothers ($P < .05$). These patterns are also reflected in breakdowns defined by parity and maternal age (Table 3). Among Whites but not Hmong, women of very high parity were less likely to enter prenatal care early. Across the maternal age groups, the same generally held true (Table 3), but not all differences were statistically significant.

TABLE 2—Cesarean Section Rate[a] by Maternal Age and Parity in Singleton Births to Hmong and White Women: Merced and San Joaquin Counties, California

Ethnicity	Age					Parity				
	<18	18–19	20–34	35–39	40+	0	1–3	4–5	6–9	10+
Hmong	3.3	5.7	3.6	5.0	4.3	11.0	3.7	1.8	3.3	6.3
White, non-Hispanic	14.8	22.2	23.8	32.6	40.0	24.4	23.8	19.1	17.1	16.7

[a]Per 100 deliveries.

TABLE 3—Percentage of Hmong and White Women with Early Prenatal Care[a] and No Prenatal Care, by Age and Parity: Merced and San Joaquin Counties, California

Ethnicity	Age									
	<18		18–19		20–34		35–39		40+	
	Early Care	No Care	Early Care	No Care	Early Care	No Care	Early Care	No Care	Early Care	No Care
Hmong	37.2	2.0	36.6	1.3	35.2	0.7	30.0	0.0	25.6	0.0
White, non-Hispanic	51.6	4.4	57.6	1.5	74.4	2.0	78.6	1.6	69.0	6.9
	Parity									
	0		1–3		4–5		6–9		10+	
	Early Care	No Care	Early Care	No Care	Early Care	No Care	Early Care	No Care	Early Care	No Care
Hmong	42.3	1.9	37.0	0.7	33.5	0.6	27.2	0.4	48.2	0.0
White, non-Hispanic	72.6	1.7	72.5	2.2	58.1	3.4	63.6	3.0	60.0	20.0

[a]Prenatal care began in months 1 to 3 of gestation.

Discussion

A previous study[12] found lower birthweight among the Chinese and Japanese in California. Our study shows that lower birthweight among Asians in California extends to the Hmong. Others[5-10] who have reported on birthweight among recent Southeast Asian immigrants to the United States have also found a lower mean birthweight among Southeast Asians compared with Whites. Information on maternal height and weight, which is not available on birth certificates, may explain the lower mean birthweight among the Hmong.

Despite the difference in mean birthweight, there is no statistically significant difference between proportions of Hmong and White births at low (< 2500 g) or very low (< 1500 g) birthweight. The similarity in the percentage of births at low or very low birthweight is most remarkable considering the high proportion of Hmong women delivering at high parity and advanced maternal age, both factors considered to be predictors of poor birthweight outcome. Moreover, although the Hmong live on public assistance, speak little English, and enter prenatal care late, they have a low absolute rate of low birthweight.

The high fertility of the Hmong in California is noteworthy. If high fertility is sustained, a study of the Hmong in California might provide information on the independent association of short interpregnancy interval and high parity with low birthweight, a controversial subject.[13-15]

Cesarean-section rates among the Hmong are strikingly lower than those among Whites in the same counties, even among women having their first child. There are anecdotal reports that Hmong refuse cesarean delivery when it is recommended, but a formal study of this has not been done. Overall complication rates are higher among White mothers than among Hmong mothers (data not shown), but the complications that are higher in Whites include dystocia, failure to progress, prior cesarean, and cephalopelvic disproportion, all of which are as much labels as diagnoses. The low cesarean-section rate among Hmong mothers provides an opportunity to determine the morbidity and mortality of low rates of cesarean section in a modern medical care setting. □

Acknowledgments

This research was supported in part by a grant from the Pacific Rim Research Program.

We thank Riley Peters at the Merced County Health Department and Ken Grodovent at the San Joaquin County Health Department for allowing us access to birth certificate data. Michael Emery did the computer programming.

References

1. Kundstadter P, ed. *Southeast Asian Tribes, Minorities and Nations.* Princeton, NJ: Princeton University Press; 1967.
2. Rumbaut R, Weeks J. Fertility and adaptation: Indochinese refugees in the United States. *Int Migration Rev.* 1986;20:428–465.
3. Strand P, Jones W. *Indochinese Refugees in America: Problems of Adaptation and Assimilation.* Durham, NC: Duke University Press; 1985.

4. Rumbaut R, Weeks J. Infant health among Indochinese refugees: patterns of infant mortality, birth weight and prenatal care in comparative perspective. *Res Sociol Health Care.* 1989;8:137–196.
5. Gann P, Nghiem L, Warner S. Pregnancy characteristics and outcomes of Cambodian refugees. *Am J Public Health.* 1989; 79:1251–1257.
6. Hopkins D, Clarke N. Indochinese refugee fertility rates and pregnancy risk factors, Oregon. *Am J Public Health.* 1983;73: 1307–1309.
7. Nelson C, Hewitt M. An Indochinese refugee population in a nurse-midwife service. *J Nurse Midwifery.* 1983;28(5):9–14.
8. Behrman R, Vaughan V, Nelson W. *Nelson Textbook of Pediatrics,* 13th ed. Philadelphia, Pa: W.B. Saunders; 1987.
9. Knuppel R, Drukker J. *High-risk Pregnancy: A Team Approach.* Philadelphia, Pa: W.B. Saunders; 1986.
10. Davis J, Goldenring J, McChesney M, Medina A. Pregnancy outcomes of Indochinese refugees, Santa Clara County, California. *Am J Public Health.* 1982;72:742–744.
11. Muecke M. Caring for Southeast Asian refugees in the USA. *Am J Public Health.* 1983;73:431–438.
12. Williams RL. Intrauterine growth curves: intra- and international comparison with different ethnic groups in California. *Prev Med.* 1975;4:163–172.
13. Klebanoff MA. Short interpregnancy interval and the risk of low birth weight. *Am J Public Health.* 1988;78:667–670.
14. Erickson JD, Bjerkedal T. Interpregnancy interval: associations with birth weight, stillbirth, and neonatal death. *J Epidemiol Community Health.* 1978;32:124–130.
15. Ferraz EM, Gray RH, Fleming PL, Maria TM. Interpregnancy interval and low birth weight: findings from a case-control study. *Am J Epidemiol.* 1988;128:1111–1116.

Call for Abstracts for National Student Research Forum

The University of Texas Medical Branch (UTMB) announces the 34th National Student Research Forum, to be held April 15 to 17, 1993, at the UTMB campus in Galveston, Texas. The Forum is intended for the competitive presentation of clinical and basic science research by medical students, graduate students, interns, and residents. It is the sole national meeting of its kind, attended by participants from throughout the United States and Canada. It is supported by grants from the American Medical Association Education and Research Foundation, Roche Laboratories, and the American Academy of Family Physicians.

Abstracts must be postmarked by no later than November 15, 1992. Further information concerning rules and awards may be obtained from the National Student Research Forum, Rm G-210, Ashbel Smith Bdg, Rte M-17, The University of Texas Medical Branch, Galveston, TX 77550; tel. (409) 772-3762.

Neonatal Tetanus in Rural Bangladesh: Risk Factors and Toxoid Efficacy

W. Gary Hlady, MD, MS, John V. Bennett, MD, Aziz R. Samadi, MD, MPH, Jahanara Begum, MBBS, MPH, Abdul Hafez, MBBS, Aminul I. Tarafdar, MBBS, MPH, and John R. Boring, PhD

ABSTRACT

Objectives. Tetanus continues to be a leading cause of neonatal death in Bangladesh as in other developing countries, yet little is known about risk factors or the efficacy of tetanus toxoid in this setting.

Methods. In May 1990, mothers of 6148 infants born alive between March 15, 1989, and March 14, 1990, in 30 rural unions of Rajshahi Division in Bangladesh were interviewed. Three surviving controls for each neonatal tetanus death were matched for sex, residence, and date of birth.

Results. Of 330 neonatal deaths, 112 met the case definition for tetanus. Risk was increased with a history of neonatal tetanus in a previous child, application of coconut oil to the vagina, and use of multiple ties on the umbilical cord. Risk was reduced by the birth attendant washing hands and using a cleaned cord-cutting tool. Risk was not reduced by a maternal history of two doses of tetanus toxoid (TT2), although estimated efficacy of TT2 was 45% (95% confidence interval = 16% to 64%). Subsequent to the survey, a reference laboratory reported no potency in three consecutive lots of tetanus vaccine from the production laboratory in Bangladesh.

Conclusions. These findings identify high-risk mothers, stress the importance of washing hands and cleaning the cord-cutting tool, and demand improved quality control of tetanus vaccine production. (*Am J Public Health.* 1992;82:1365–1369)

Introduction

Neonatal tetanus continues to be a major cause of death in developing countries, accounting for more than 654 000 deaths in 1990,[1] yet few analytic studies have been done. The value of such studies was most recently demonstrated by a report from Pakistan, which suggested that the risk of neonatal tetanus could be significantly reduced by eliminating the application of ghee (clarified butter) to the umbilical stump during the first several days of life—a common practice in that country and a previously unidentified risk factor.[2] Neonatal tetanus can be totally prevented both by two injections of tetanus toxoid vaccine (TT2) to the mother, with the second dose at least 4 weeks after the first and at least 30 days prior to delivery, and by sterile cord care. But vaccination coverage of pregnant women in most developing countries remains low, and unsafe birth practices persist.[1]

Bangladesh has one of the highest rates of neonatal tetanus in the world.[1] The government has addressed this problem through a vigorous program of vaccination, training of traditional birth attendants, and promotion of safe-delivery kits. In addition, a national study was begun in February 1990 to identify risk factors for neonatal tetanus. This study, which was patterned after the study in Pakistan, was conducted by Global 2000, Inc, in collaboration with the Ministry of Health and Family Planning, and the Task Force for Child Survival and Development. We report here the results of the Rajshahi Division survey, with emphasis on the findings not only of potential risk factors but also of the low efficacy of tetanus toxoid, apparently due to the use of low-potency vaccine.

Methods

Study Population

Bangladesh is divided into four administrative divisions of approximately equal populations. The survey reported here was of Rajshahi Division, occupying the northwest quadrant of the country with a 1981 rural population of 20 009 351.[3] Cluster sampling was done according to World Health Organization (WHO) guidelines.[4] Thirty rural unions (average population 18 000) were systematically selected from the total of 1089 listed in the 1981 census by using a random start and a sampling interval equal to the cumulative population divided by 30, making the probability of selection proportionate to population size. Urban areas were excluded.

Each union is divided into three wards of approximately equal populations, and each ward is divided into a variable number of *mauzas* (average population 1400). The starting point for the survey in each selected union was the first

W. Gary Hlady is with the Epidemiology Program of the Department of Health and Rehabilitative Services in Tallahassee, Fla; John V. Bennett is with the Task Force for Child Survival in Atlanta, Ga; Aziz R. Samadi and John R. Boring are with the Epidemiology Division of Emory University's School of Public Health in Atlanta; and Jahanara Begum, Abdul Hafez, and Aminul I. Tarafdar are with the National Institute of Preventive and Social Medicine in Mohakhali, Dhaka, Bangladesh.

Requests for reprints should be sent to Dr. W. Gary Hlady, Epidemiology Program, Department of Health and Rehabilitative Services, 1317 Winewood Blvd, Bldg 2, Rm 247, Tallahassee, FL 32301.

This paper was submitted to the Journal December 4, 1991, and accepted with revisions May 22, 1992.

TABLE 1—Circumstances of Delivery for Neonatal Tetanus Cases (n = 112) and Matched Controls (n = 336)

	Cases		Controls	
	No.	%	No.	%
Place of delivery				
In hospital or clinic	0	0	6	1.8
At home	112	100	330	98.2
Birth attendant				
Relative	63	56.3	182	54.2
Unrelated traditional birth attendant	23	20.5	84	25.0
Health care worker	0	0	8	2.4
Self (delivered without assistance)	26	23.2	62	18.4
Delivery practices[a]				
Attendant washed hands	45	40.2	182	54.2
Perineum washed	41	36.6	146	43.5
Mustard oil applied to vagina	22	19.6	69	20.5
Coconut oil applied to vagina	9	8.0	12	3.6
Safe-delivery kit used	2	1.8	7	2.1
Cord-cutting tool cleaned before use	55	49.1	231	68.8
Multiple cord ties used	41	36.6	89	26.5
Mother's vaccination status at delivery				
History of TT2[b]	33	29.5	122	36.3
No history of TT2	79	70.5	214	63.7

[a]Figures exceed total (100%) because categories are not exclusive.
[b]Two doses of tetanus toxoid at least 4 weeks apart, with the second dose at least 30 days before delivery.

household encountered in a *mauza* selected by use of a random number table. The survey then proceeded from house to house within the same ward until a target of 67 mothers of eligible infants were interviewed. An eligible infant was one who was born alive between March 15, 1989, and March 14, 1990. The survey continued then from the nearest house in the next nearest ward and so on until all three wards were surveyed, for a total target of 201 interviews in each selected union. Interviews completed in excess of the target (usually 3 or 4 per union) were also included in the analysis.

Interviews were conducted by 10 teams of six female interviewers using a standardized questionnaire. Interviewers received 6 days of instruction by Global 2000 staff and faculty of the National Institute of Preventive and Social Medicine, and 4 days of closely supervised field testing. The survey was completed between May 3 and May 28, 1990.

Case Definition

This was a case-control study. A case was an eligible infant who was normal at birth but who died at 3 to 30 days of age following illness, with generalized spasms and at least three of the following four signs: trouble opening mouth, cessation of sucking, clenched hands, and boardlike rigidity. Infants who were seen at a health care facility and given a diagnosis other than tetanus (as reported by the mother) were excluded as cases.

Control Selection

Three living infants—matched for sex, residence (union), and date of birth to control for the effects of circumcision, regional variations in birth practices, and recall—were selected as controls for each case. When more than three infants were eligible as controls, priority for selection was given to those with birth dates closest to the case's and to those who resided in the same ward, in that order. Control selection differed from that used in the Pakistan study in that the mothers' vaccination status with tetanus toxoid was disregarded in selecting control infants.

Data Analysis

Data were double entered into an Epi Info, Version 5, data file and were analyzed using the same software package.[5] Univariate analysis of matched data used the same calculations as stratified Mantel-Haenszel analysis to calculate matched odds ratios, with each stratum consisting of a single matched case-control group. The 95% confidence intervals (CIs) for matched odds ratios were calculated using standard methods,[6] and all probabilities were based on two-tailed tests.

Multivariate analysis was performed using EGRET software.[7] Data were defined by the DEF program for conditional logistic regression and analyzed by the PECAN program.[7] The log likelihood ratio statistics, degrees of freedom, and probability statistics for the model, and the 95% CIs and *P* values for each variable, were calculated by the program. Interaction was examined by evaluating the *P* value for main and product terms as well as significant differences in likelihood ratio statistics between noninteractive and interactive models. Multicolinearity among covariables was examined by modeling and correlation procedure.

Vaccine efficacy based on the entire study population was calculated using the following formula: $(Iu-Iv)/Iu$, where Iu is the incidence of neonatal tetanus in infants with no maternal history of tetanus toxoid vaccine prior to delivery and Iv is the incidence of neonatal tetanus in infants with a maternal history of TT2. This formula is equal to $1-(Iv/Iu)$, or 1–(risk ratio), and 95% CIs for the risk ratio were used to calculate confidence intervals for efficacy.[8] Because vaccine efficacy was not an anticipated issue, incidence rates could not be weighted to account for differences in sample union populations. However, the difference between weighted and unweighted rates using this sampling procedure (the EPI method) is considered small and acceptable.[9]

Results

Descriptive Statistics

A total of 43 791 households were surveyed, and the mothers of 6148 eligible infants were interviewed. The circumstances of delivery are summarized in Table 1. There were 330 neonatal deaths (occurring up to 30 days of age). One hundred twelve neonatal deaths met the case definition for tetanus, resulting in a neonatal tetanus death rate of 18.2 per 1000 live births in our study population. Three infants who otherwise met the case definition were excluded because they had been seen at a health care facility and the mothers reported a cause of death other than tetanus (two "evil spirits," one "milk disease"). Only 21 (18.8%) of the cases were seen at a health care facility.

Male cases tended to be seen more often than female cases at a health care facility (23% vs 15%, $P = .3$), but the proportion of males among cases (50.9%) did not differ from that among all live births (51.1%).

None of the male infants was circumcised before 30 days of age.

The dates of birth for cases and for all live births are shown in Figure 1. The cyclic pattern of live births is identical with that observed prospectively in Matlab, Bangladesh.[10] The July through November peak in neonatal tetanus deaths has also been previously observed in India.[11–13] The mean age at death for cases was 10.7 days (SD = 6.3); the median age was 8 days.

Application of substances to the vagina and to the cord stump at the time of delivery were common practices and usually involved the use of mustard oil, which was heated before use. Mustard oil was also commonly applied several times to the cord stump during the first few days of life. Ghee was never applied to either the vagina or the cord stump.[2] Cow dung was applied to the cord stumps of only three infants on the day of delivery, one of whom died of tetanus. Cow dung was also applied to the cord stumps of eight infants only on days subsequent to birth, and none of those infants died of tetanus.

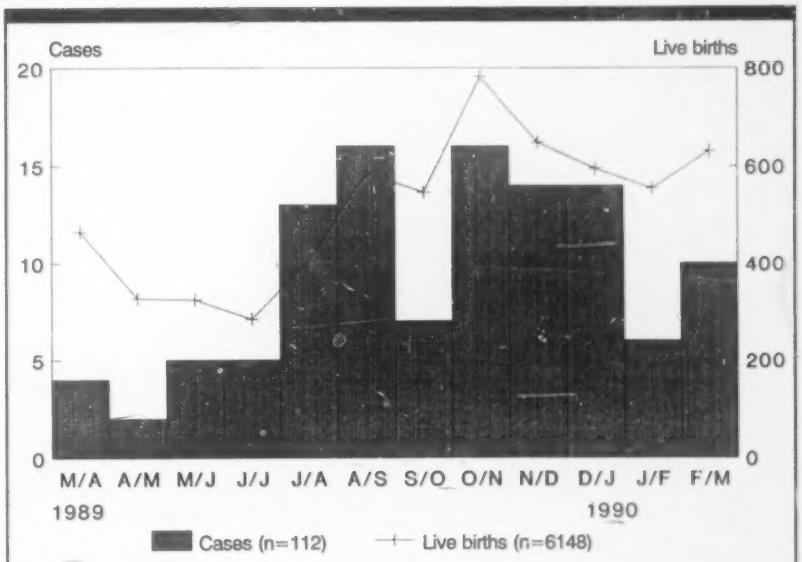


FIGURE 1—Month (English equivalent of Bengali month) of birth for neonatal tetanus cases and all live births.

TABLE 2—Multivariate Analysis[a] of Risk Factors for Neonatal Tetanus

	Coefficient	Standard Error	Odds Ratio	95% CI	*P*
History of TT2[b]	–0.2712	.267	0.76	0.45, 1.29	.3
Hand washing	–0.7150	.257	0.49	0.30, 0.81	.005
Cleaned cutting tool	–0.9234	.258	0.40	0.24, 0.66	<.001
Vaginal coconut oil	1.150	.526	3.16	1.13, 8.86	.03
Multiple cord ties	0.6919	.276	2.00	1.16, 3.43	.01

Note. Likelihood ratio statistic on 5 *df* = 34.711, *P* < .001.
[a]Conditional logistic regression by PECAN procedure.[7]
[b]Two doses of tetanus toxoid at least 4 weeks apart, with the second dose at least 30 days before delivery.

Risk Factors

The results of a multivariate analysis of risk factors for neonatal tetanus are shown in Table 2. The striking finding was that a maternal history of TT2 failed to be protective. However, hand washing and the use of a cleaned tool to cut the umbilical cord, regardless of what tool was used, were both associated with decreased risk. The risk increased when coconut oil was applied to the vagina and when the umbilical cord was tied in more than one place.

A history of neonatal tetanus in a previous child was a significant risk factor; its odds ratio, obtained by conditional logistic regression as a single variable in the model, was 3.63 (95% CI = 1.30, 10.14; $P = .01$), but it was excluded from the final model because of colinearity with hand washing. This statistical association was also confirmed by stratified Mantel-Haenszel analysis. Such a history was present in 27.7% (31/112) of cases and in only 7.7% (26/336) of controls. Of previous children born to case mothers, 14.3% (39/273) were reported to have died of neonatal tetanus, compared with 4.5% (33/740) of previous children born to control mothers ($\chi^2 = 29.14$, $P < 10^{-6}$). The mean total number of live births was 3.4 (SD = 2.6) for mothers of cases and 3.2 (SD = 2.2) for mothers of controls. Similarly, mean maternal age was 26.1 years (SD = 6.3) for cases and 25.5 years (SD = 5.7) for controls.

Other factors of interest, including use of a safe-delivery kit, washing of the perineum, application of substances to the umbilical cord, and the presence of a trained birth attendant (8 cases, 35 controls), were not associated with risk and thus were not included in the final regression model.

Efficacy of TT2

The maternal vaccination status at the time of delivery for neonatal tetanus cases and all other live births in our study population is shown in Table 3. The efficacy of a maternal history of TT2, based on these figures, was 45% (95% CI = 16%, 64%). The efficacy of a maternal history of TT2, based on the odds ratio obtained by multivariate analysis of the case-control data, was 24% (95% CI = −29%, 55%). Although less precise, the latter efficacy estimate may be more valid as it is adjusted for the effects of other risk factors. However, both estimates are far below the level expected for a potent vaccine. Verification of vaccination by medical records was difficult to achieve because the medical records were kept with mothers, and only 48.3% of all mothers with a history of TT2 (1209/2505) possessed a medical record at the time of interview.

Discussion

The accuracy of our case ascertainment is supported by the fact that the rate of neonatal tetanus death observed in our study population is similar to that reported by others.[14–16] Had we used a less restrictive case definition of only a history of generalized spasms preceding death during the first 30 days of life, as has been used elsewhere,[17] the number of neonatal tetanus deaths in our sample would have increased to 259 (42.1 per 1000 live births). A similar increase would have resulted

TABLE 3—Maternal Vaccination Status at Time of Delivery

	Neonatal Tetanus Cases (n = 112)		No Neonatal Tetanus (n = 6036)		All live births (n = 6148)	
Vaccination Status	No.	%	No.	%	No.	%
None[a]	56	50.0	2278	37.7	2334	38.0
History of TT1[b]	23	20.5	1286	21.3	1309	21.3
History of TT2[c]	33	29.5	2472	41.0	2505	40.7

[a]Never vaccinated.
[b]Received one dose of tetanus toxoid at least 30 days before delivery. (May have received a second dose less than 30 days before delivery.)
[c]Received two doses of tetanus toxoid, with second dose at least 30 days before delivery.

had we relied only on the mothers' report of cause of death. We chose the more specific case definition to maximize the accuracy of diagnosis for the purpose of identifying risk factors, but this may have also resulted in an underestimate of the incidence of neonatal tetanus death.

Our observation that tetanus toxoid failed to provide the expected high level of protection was not the first to indicate that there may be a problem with the Bangladesh vaccine. Anecdotal reports from physicians had already described apparent vaccine failures; and a 1989 WHO survey of more than 4000 live births, equally divided between areas with high vaccine coverage and areas with low vaccine coverage, found almost twice the expected rate of neonatal tetanus death in the high coverage area.[18]

Subsequent to our findings, a WHO reference laboratory reported no potency in samples from three consecutive lots of tetanus toxoid produced in Bangladesh in July 1990 (J. Milstien, WHO, personal communication, October 17, 1990), and a serosurvey of 100 women vaccinated in September and October 1990 with two doses of tetanus toxoid vaccine found that only 31% had developed protective levels of antibody (N. Abiprojo, UNICEF, personal communication, July 29, 1991).

Since 1983, routine potency testing of tetanus toxoid in Bangladesh has been performed in the same facility by the same staff that produced all the tetanus toxoid vaccine for the national immunization program. There is no national control authority in Bangladesh to certify vaccine safety or potency. Currently, tetanus toxoid production in Bangladesh has been suspended, and the immunization program has continued with the use of imported vaccine.

It has not yet been determined when the problem with vaccine quality control began. Documented vaccination histories from Rajshahi suggest that the problem existed with vaccine in the field at least as early as February 1989. With a 1- to 3-month-long distribution "pipeline," this would include all vaccine produced since January 1989—approximately two thirds of all tetanus vaccine ever produced in Bangladesh—with more than 40 000 000 doses administered. It is possible that the problem was intermittent, but this cannot be determined from available information as no records are kept of the distribution of vaccine by lot number.

Nevertheless, a potent preparation of tetanus toxoid is one of the safest, most effective vaccines available and is the primary element of any neonatal tetanus control program.[1,14] Bangladesh has achieved a remarkably high level of vaccine distribution. If immunization activities are sustained, corresponding decreases in tetanus mortality can be expected with the introduction of a potent vaccine. The very high risk associated with a maternal history of previous neonatal tetanus death, which was initially observed elsewhere[2] and was confirmed in this study, should make this group of women a particularly high priority for immunization.

Failure of the tetanus vaccine in Bangladesh, however, must also serve as a reminder that such high-technology interventions are not always effectively implemented and should not be relied upon solely, especially where simple measures such as hand washing and cleaning of the cord-cutting tool have proven effective and reliable.

The use of a trained traditional birth attendant was not observed to be protective, possibly because of inaccuracies in both the mothers' ascertainment of attendant training and the low number of trained attendants involved. For these reasons and because certain behaviors of the birth attendants were significantly related to decreased risk, our results should not discourage the training of traditional birth attendants as a control measure for neonatal tetanus. Similarly, the low number of safe-delivery kits used did not allow a meaningful evaluation of their effect.

The increased risk associated with the use of multiple cord ties has not been observed elsewhere and awaits further clarification. However, the increased risk associated with the vaginal use of coconut oil is an important finding and is in line with the risk associated with the vaginal use of ghee in Pakistan (H. Traverso, personal communication, December 18, 1990). The mechanism by which the vaginal use of coconut oil increases the risk of neonatal tetanus is unknown, but the increased risk could be due to contamination of the oil by the uncleaned hands of the birth attendant.

Unfortunately, control of neonatal tetanus in Bangladesh will be delayed due to the failure to maintain vaccine potency. Given that fewer than one in five neonatal tetanus cases had contact with the health care system and that the problem with the tetanus vaccine was not brought to light earlier, the need for improved surveillance to evaluate program effectiveness and to monitor vaccine quality and distribution in the field is indicated. Such improvement should include the creation of a separate quality control laboratory and a national control authority. More work is needed to determine how surveillance might best be structured and what thresholds for action are appropriate. □

Acknowledgments

Funding for this work was provided by the BCC Foundation, Bangladesh.

The authors are grateful for technical assistance and administrative support provided in Bangladesh by Prof. Md. Khabiruddin and Prof. A. M. Das of the National Institute of Preventive and Social Medicine; Dr. A. K. M. Lutfar Rahman Talukdar, director, Expanded Programme on Immunization; Prof. Farida Huq, director, Institute of Public Health; Mr. S. Y. Khan Mojlish, Ministry of Health and Family Planning; and other members of the Neonatal Tetanus Control Project Coordinating Committee.

References

1. Gasse F. Neonatal tetanus elimination initiative: progress report and recommendations. Geneva, Switzerland: World Health Organization; 1990. WHO Document EPI/MCH/NNT/GEN/90.1.
2. Traverso HP, Bennett JV, Kahn AJ, et al. Ghee applications to the umbilical cord: a risk factor for neonatal tetanus. *Lancet.* 1989;1:486–488.
3. *1989 Statistical Yearbook of Bangladesh.* Dhaka, Bangladesh: Bangladesh Bureau of Statistics; 1989.

4. Galazka A, Stroh G. Neonatal tetanus guidelines on the community-based survey on neonatal tetanus mortality. Geneva, Switzerland: World Health Organization; 1986. WHO Document EPI/GEN/86/8.
5. Dean AD, Dean JA, Burton AH, Dicker RC. *Epi Info, Version 5: A Word Processing, Database and Statistics Program for Epidemiology on Microcomputers.* Stone Mountain, Ga: USD, Inc, 1990.
6. Robins JM, Greenland S, Breslow NE. A general estimator for the variance of the Mantel-Haenszel odds ratio. *Am J Epidemiol.* 1986;124:719–723.
7. *EGRET Manual Addendum, Revision 1.* Seattle, Wash: Statistics and Epidemiology Research Corporation; 1990.
8. Clemens JD, Shapiro ED. Resolving the pneumococcal vaccine controversy: are there alternatives to randomized clinical trials? *Rev Infect Dis.* 1984;6:589–600.
9. Lemeshow S, Stroh G. *Sampling Techniques for Evaluating Health Parameters in Developing Countries.* Washington, DC: National Academy Press; 1988:11. Monograph.
10. Chowdhury MK, Karim MR, Mostafa G, Sarder AM, D'Sousa S. *Demographic Surveillance System—Matlab.* Vol 11. Dhaka, Bangladesh: International Centre for Diarrhoeal Disease Research; 1983. B Scientific Report No. 59.
11. Jaffari SMH, Pandit MM, Ismail M. Neonatal tetanus in Hyderabad. *Indian Pediatr.* 1966;3:177–182.
12. Nigam PBM, Goyal BM, Tandon VK. Tetanus neonatorum—a clinical study of 72 cases. *Indian J Pediatrics.* 1974;41:135–143.
13. Laha PN, Vaishya PD. A study of the incidence of 1000 cases of tetanus in Gwalior and its neighboring districts. *J Indian Med Assoc.* 1965;44:422–436.
14. Bhatia S. Patterns and causes of neonatal and postneonatal mortality in rural Bangladesh. *Stud Fam Plann.* 1989;20:136–146.
15. Islam MS, Rahaman MM, Aziz KMS, Munshi MH, Rahaman M, Patwari Y. Birth care practice and neonatal tetanus in a rural area of Bangladesh. *J Trop Pediatr.* 1982;28:299–302.
16. Chen LC, Rahman M, Sarder AM. Epidemiology and causes of death in a rural area of Bangladesh. *Int J Epidemiol.* 1980;9:25–33.
17. Kalter HD, Gray RH, Black RE, Gultiano SA. Validation of post mortem interviews to ascertain selected causes of death in children. *Int J Epidemiol.* 1990;19:380–386.
18. Basu RN. Neonatal tetanus in Bangladesh. *In Touch* (Voluntary Health Services Society newsletter, Dhaka, Bangladesh). 1989;13:3–5.

Improving Estimates of HIV-1 Seroprevalence among Childbearing Women: Use of Smaller Blood Spots

Neil J. Hoxie, MS, James M. Vergeront, MD, John R. Pfister, BS, Gary L. Hoffman, BS, Pamela A. Markwardt-Elmer, BS, and Jeffrey P. Davis, MD

ABSTRACT

Objectives. **Nationwide, human immunodeficiency virus type 1 (HIV-1) seroprevalence surveys using dried neonatal blood specimens are critical to estimating HIV-1 seroprevalence among childbearing women. However, the noninclusion of blood specimens deemed "quantity not sufficient" (QNS) for HIV-1 antibody testing potentially introduces bias. In Wisconsin beginning in 1990, we modified the survey protocol to reduce QNS rates and assess bias introduced by QNS specimens.**

Methods. **The HIV-1 antibody assay was modified to use four 1/8-in blood spots when a single 1/4-in blood spot could not be obtained. Both methods obtain identical blood volumes for testing.**

Results. **During a 27-month period, 7396 (4.8%) of 154 683 specimens were deemed QNS using 1/4-in blood spots. Of these, 6590 (89%) were of sufficient quantity to be tested using four 1/8-in blood spots; 6 (0.09%) specimens tested with 1/8-in blood spots were HIV-1 Western blot assay positive compared with 44 (0.03%) of 147 287 1/4-in specimens (odds ratio = 3.0; 95% confidence interval = 1.2, 7.4).**

Conclusions. **Because noninclusion of QNS specimens potentially introduces bias, incorporating the results of HIV-1 antibody testing of QNS specimens using four 1/8-in blood spots can improve the accuracy of HIV-1 seroprevalence estimates in these serologic surveys. (*Am J Public Health.* 1992;82:1370–1373)**

Introduction

In the United States, surveillance of acquired immunodeficiency syndrome (AIDS) has demonstrated increasing morbidity among both women and children of women with human immunodeficiency virus type 1 (HIV-1) infection.[1,2] Responding to the epidemic of HIV-1 infection among women and children requires accurate information on the prevalence and distribution of this infection within these populations. To provide this information, the Centers for Disease Control (CDC) and the National Institute for Child Health and Development have developed a standard survey protocol for estimating HIV-1 seroprevalence among childbearing women. This protocol uses dried neonatal blood specimens collected on filter paper as part of newborn screening programs for blinded HIV-1 antibody testing. By March 1990, surveys using this protocol were ongoing in 43 states.[3] One such survey has been conducted since 1989 by the AIDS/HIV Program of the Wisconsin Division of Health and the Wisconsin State Laboratory of Hygiene in collaboration with the CDC.

The standard protocol for this survey specifies that 100% of blood specimens received by newborn screening programs during specified periods be tested for HIV-1 antibody.[4] However, attaining this goal is difficult. After mandated newborn screening tests are completed, some specimens are inevitably designated as being of a "quantity not sufficient" (QNS) for HIV-1 antibody testing. The inability to test QNS specimens can result in sampling bias if the prevalence of HIV-1 antibody among QNS specimens is different from that among specimens available for testing. If sampling bias is operative, the accuracy of HIV-1 seroprevalence estimates may be adversely affected.

To investigate whether an adverse effect has occurred in Wisconsin, beginning in 1990 dried blood specimens designated as QNS using standard specimen preparation methods were tested for HIV-1 antibodies using an alternate method. Objectives of the study included determining whether use of the alternate specimen preparation method could decrease the number of QNS specimens and whether the noninclusion of QNS specimens does bias estimates of HIV-1 seroprevalence among childbearing women.

Methods

The standard CDC protocol for blinded HIV-1 seroprevalence surveys using dried neonatal blood specimens has been previously described.[3] Briefly, as relates to this study, capillary blood specimens from neonates born in Wisconsin are collected on a standardized filter paper form, air dried, and sent to the Newborn Screening Unit of the Wisconsin State Laboratory of Hygiene. Duplicate and unsatisfactory specimens are excluded from the survey; only one specimen obtained

Neil J. Hoxie, James M. Vergeront, and Jeffrey P. Davis are with the Bureau of Public Health, Wisconsin Division of Health, in Madison, Wisconsin. John R. Pfister and Pamela A. Markwardt-Elmer are with the Retrovirus Immunology Unit, and Gary L. Hoffman is with the Newborn Screening Unit, of the Wisconsin State Laboratory of Hygiene in Madison.

Requests for reprints should be sent to Neil J. Hoxie, MS, 1414 E Washington Ave, Room 96, Madison, WI 53703.

This paper was submitted to the Journal September 23, 1991, and accepted with revisions May 5, 1992.

from children of multiple births is included. The remaining specimens are eligible for inclusion in this survey once routine, mandated newborn screening tests are completed and personal identifiers have been removed.

Specimens are prepared for HIV-1 antibody testing by punching a blood spot from an area of filter paper judged to be adequately saturated with blood. In Wisconsin, blood spots are routinely obtained using an automated 1/4-in-diameter circular punch (Fundamental Products, North Hollywood, Calif). Use of 1/4-in blood spots is specified in the CDC protocol and is referred to herein as the standard method. If the surface area of blood-saturated filter paper is insufficient to allow a standard punch, the specimen is considered QNS. In this study, if a specimen was deemed QNS, a manual 1/8-in punch (McGill Metal Products, Marengo, Ill) was used to obtain four 1/8-in-diameter circular spots from specimens with sufficient blood-saturated surface area. This is referred to herein as the alternate method. Both the standard and the alternate punching methods result in the collection of an identical surface area of blood-saturated filter paper and thus an identical volume of blood for HIV-1 antibody testing.

Standard and alternate specimens were tested by the Wisconsin State Laboratory of Hygiene's Retrovirus Immunology Unit for HIV-1 antibody using identical procedures, which have been previously described.[5,6] Specimens were tested by enzyme-linked immunosorbant assay (ELISA) (Abbott Laboratories, Chicago, Ill) and were retested in duplicate if reactive. If one or more of the repeat assays were reactive, a Western blot assay (Organon Tecknika, Durham, NC) was performed on the same specimen. Specimens were considered HIV-1 antibody positive if at least two bands from among p24, gp41, and gp120/160 were present on the Western blot.[7]

Parallel testing, using 20 samples each of HIV-1 antibody negative and high and low HIV-1 antibody positive dried blood controls provided by the CDC, was conducted to determine if HIV-1 ELISA results for standard and alternate blood spots were comparable. Comparability was also assessed using a panel of 31 HIV-1 antibody positive and 30 HIV-1 antibody negative dried blood specimens prepared by the Retrovirus Immunology Unit. In addition, a CDC proficiency panel consisting of 12 specimens of unknown HIV-1 reactivity was analyzed by both HIV-1 ELISA and Western blot assay using 1/8-in blood spots, and the observed results compared with those expected by CDC referees. Only after the test methods were judged comparable were the small spot specimens included in the survey.

TABLE 1—Results of HIV-1 Antibody Testing of Dried Blood Spots of Filter Paper Impregnated with Neonatal Blood, Wisconsin, January 1990 through March 1992

	Standard (1/4-in) Punches[a]			Alternate (1/8-in) Punches[b]					
	Tested	HIV+	%+	Tested	HIV+	%+	Odds Ratio	95% CI	*P*
Maternal race									
White	126 024	17	0.01	5311	2	0.04	2.6[c]	1.0, 6.3[c]	.06[c]
Minority[d]	21 258	27	0.13	1279	4	0.31			
Maternal age									
<20 y	15 319	8	0.05	793	0	0.0	3.0[c]	1.2, 7.4[c]	.02[c]
≥20 y	131 949	36	0.03	5796	6	0.10			
Total	147 287	44	0.03	6590	6	0.09	3.0	1.2, 7.4	.02

[a]Specimens were obtained with 1/4-in punch as specified in the standard Centers for Disease Control protocol.
[b]Specimens were deemed "quantity not sufficient" (QNS) by standard methods and were obtained using four 1/8-in punches.
[c]Statistical analysis compared alternate punched specimens with standard punched specimens, adjusted using the Mantel-Haenszel method.
[d]Minority category includes all races other than White; Hispanic may be classified as either White or minority.

Data were analyzed using Epi Info, Version 5.0, software (CDC, Atlanta, Ga). Statistical tests included odds ratios (ORs), Cornfield's 95% confidence intervals (CIs) of odds ratios, and tests of significance using either the Yates Corrected χ^2 test if all individual cells were greater than or equal to 5, or Fisher's Exact Test if any individual cell was less than 5. Adjusted analysis was performed using the Mantel-Haenszel method.[8] Results of parallel testing of CDC blood spot controls were compared using a *t* test; absorbance ratios of comparison panel specimens were compared using a paired *t* test.

The HIV-1 seroprevalence rate is the proportion of specimens tested that was HIV-1 Western blot positive. The QNS rate is the proportion of the total sample determined to be QNS by standard punching methods. The salvage rate is the proportion of QNS specimens that could be sampled using multiple 1/8-in punches. Bias factor is defined as the HIV-1 seroprevalence rate among specimens tested using 1/8-in spots divided by the rate among 1/4-in specimens. If "Sero A" refers to the combined total seroprevalence of 1/4-in and 1/8-in blood spots and "Sero B" to the seroprevalence of only 1/4-in spots, relative error (RE) can be defined as ([Sero A − Sero B]/ Sero A) × 100. RE can also be expressed as a function of the QNS rate and the bias factor: RE = QNS rate × (bias factor − 1) × 100.

Results

From January 1990 through March 1992, a total of 154 683 specimens were eligible for inclusion in this survey. Of these, 7396 (4.8%) were deemed of insufficient quantity to test for HIV-1 antibody using a standard 1/4-in punch. The QNS rate was higher among specimens collected from minority neonates (1457/22 715, 6.4%) than among specimens from White neonates (5937/131 961, 4.5%) ($P < .001$, Yates Corrected χ^2), and higher among specimens collected from neonates with teenage mothers (885/16 204, 5.5%) than among specimens from neonates with older mothers (6508/138 457, 4.7%) ($P < .001$, Yates Corrected χ^2) (Table 1).

Sufficient acceptable specimen was available to obtain four 1/8-in blood spots from 6590 (89%) of the 7396 QNS specimens. The salvage rate was similar among specimens from White neonates (5311/5937, 89%) and minority neonates (1279/1457, 88%), as it was among specimens from neonates with teenage mothers (793/885, 90%) and neonates with older mothers (5796/6508, 89%).

The HIV-1 seroprevalence rate among specimens punched using the alternate method was threefold greater (.09% vs .03%) than that among specimens punched using the standard method

(odds ratio [OR] = 3.0; 95% confidence interval [CI] = 1.2, 7.4; $P = .02$, Fisher's Exact Test) (Table 1). After adjusting for race, the point estimate of the adjusted OR (2.6) was slightly less than that of the crude OR (3.0). Adjusting for maternal age resulted in an OR identical to the crude OR. Using only 1/4-in blood spots, the estimate of HIV-1 seroprevalence among childbearing women is 0.0299%; when 1/4-in and 1/8-in blood spot test results are combined, however, the HIV-1 seroprevalence estimate increases 8.7% to 0.0325%.

The results of 20 HIV-1 ELISA procedures performed on the negative CDC control specimen demonstrated a mean absorbance ratio of 0.32 (range = .22 to .43) for 1/8-in blood spots compared with 0.37 (range = .27 to .50) for 1/4-in blood spots ($P < .01$). Twenty HIV-1 ELISA procedures performed on the low positive CDC control specimen produced a mean absorbance ratio of 2.68 (range = 2.36 to 3.21) for 1/8-in blood spots and 2.46 (range = 1.62 to 3.30) for 1/4-in blood spots ($P = .07$). Twenty HIV-1 ELISA procedures performed on the high positive CDC quality control specimen resulted in a mean absorbance ratio of 8.29 (range = 7.25 to 10.04) for 1/8-in blood spots and 7.26 (range = 5.36 to 8.59) for 1/4-in blood spots ($P < .01$).

For both 1/8-in and 1/4-in negative control blood spots, HIV-1 ELISA absorbance ratios were all well below the testing laboratory's cutoff absorbance ratio (< 1.0) for nonreactive HIV-1 ELISA; thus, no misclassification of HIV-1 serostatus occurred among negative specimens. The absorbance ratios of all low and high positive CDC quality control specimens tested using 1/8-in and 1/4-in blood spots were all well above the cutoff absorbance (≥ 1.0) for reactive HIV-1 ELISA; thus, no misclassification of HIV-1 serostatus occurred among positive specimens.

Comparison of paired HIV-1 ELISA test results (1/8-in blood spots vs 1/4-in blood spots) of the HIV-1 negative and the HIV-1 positive comparison panels prepared by the testing laboratory demonstrated no discordant classification of HIV-1 serostatus. Paired analysis of the 30 HIV-1 negative panel specimens showed the mean of the absorbance ratio differences to be -.022 (SE = .023; $P = .33$, paired t test). Paired analysis of the 31 HIV-1 positive panel showed the mean of the absorbance ratio differences to be -.082 (SE = .10; $P = .41$, paired t test).

HIV-1 ELISA testing of the 12 CDC proficiency panel specimens using 1/8-in blood spots resulted in the correct HIV-1 classification for each specimen. In this panel, the expected banding patterns of the Western blot assays were demonstrated on each of the six specimens derived from the 1/8-in blood spots that yielded reactive HIV-1 ELISA results.

Discussion

Our findings demonstrate that, in HIV-1 seroprevalence surveys that use dried neonatal blood specimens, QNS rates can be reduced by using multiple 1/8-in blood spots to analyze specimens that are of insufficient quantity to test using a standard 1/4-in punch. Use of this technique provided an opportunity to analyze 89% of specimens that would otherwise have been unavailable for HIV-1 antibody testing.

Our comparison studies indicate that, although multiple 1/8-in blood spots may produce different absorbance ratios on HIV-1 ELISA, HIV-1 antibody testing of these small spots results in an HIV-1 antibody status classification identical to that obtained testing standard-size blood spots. This suggests that multiple 1/8-in blood spots can be reliably tested for HIV-1 antibody with the Abbott assay using the same testing methods as for 1/4-in spots. Laboratories that use other assays should conduct parallel testing to confirm the comparability of results.

The presence of QNS specimens would have no adverse effect on overall estimates of HIV-1 seroprevalence if the seroprevalence of QNS specimens were similar to that of specimens tested using standard punches. However, an important finding of this study is that the prevalence of HIV-1 antibody among QNS specimens was threefold greater than that among specimens available by standard sampling methods; this observed difference was independent of maternal race and age. In our study, use of multiple 1/8-in blood spots increased the number of HIV-1 seropositive specimens by 14% and the estimate of HIV-1 seroprevalence among childbearing women by nearly 9%. Thus, the occurrence QNS specimens *can* bias HIV-1 seroprevalence estimates.

Reasons for the occurrence of QNS specimens probably differ from state to state. They could result from inadequate specimen collection although, in Wisconsin, poorly collected specimens are not accepted for newborn screening and a second specimen is requested. It has been the experience of the Wisconsin Newborn Screening Unit, however, that many QNS specimens have had additional newborn screening tests performed by necessity on the initial sample (Gary L. Hoffman, unpublished data). This additional testing results when the values for initial newborn screening tests fall outside of established limits, in which case it is laboratory policy to repeat those tests. Additional testing results in an increased consumption of the blood specimen, which may result in an insufficient quantity of dried blood to allow HIV-1 antibody testing.

The reasons for our finding of higher HIV-1 seroprevalence among QNS specimens are unclear. It is possible, however, that the requirement for additional testing, which leads to QNS specimens, may be associated with—among other factors—low birthweight. It has been observed that low-birthweight infants are more likely to have newborn screening test results, particularly those for thyrotropin and thyroxine, outside the acceptable limits established by the newborn screening laboratory. Additionally, low-birthweight infants are more likely to be treated with antibiotics, aminosin, or total parenteral nutrition compared with normal birthweight infants. Because these treatments can interfere with some newborn screening assays, spurious laboratory values can result, increasing the need for additional testing.

A correlation of QNS specimens with low birthweight could help explain the association of QNS specimens with teenage mothers and minority status. In Wisconsin, low birthweight is associated with both of these characteristics[9]; in New York City, it was found to be closely correlated with newborn HIV seropositivity.[10] And while the etiology of low birthweight is multifactorial, in some cases it is associated with injection drug use,[11–13] which in turn is strongly associated with HIV-1 infection among women and their children.[14,15] Thus, although this conclusion is speculative, the higher HIV-1 seroprevalence among specimens obtained using 1/8-in punches could be a result of an increased prevalence of maternal intravenous drug use among neonates with QNS neonatal blood specimens.

In certain situations, larger relative errors resulting from QNS bias are plausible. As shown in Figure 1, relative error introduced by this effect is a function of both the QNS rate and the bias factor. The relative error produced could be of public health significance, even at moderate QNS rates. For example, with a QNS rate of 10%, HIV-1 seroprevalence could be underestimated by 10% if the bias factor is equal to 2 and by 40% if the bias factor is

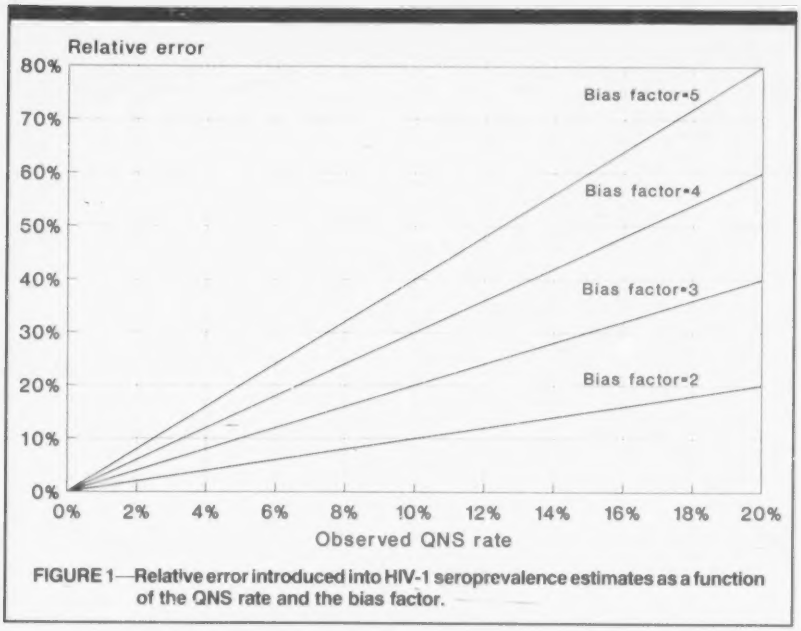


FIGURE 1—Relative error introduced into HIV-1 seroprevalence estimates as a function of the QNS rate and the bias factor.

equal to 5. Thus, minimizing QNS rates is crucial to the reduction of relative error resulting from sampling that is biased by QNS specimens.

The alternate method we describe is only marginally more labor-intensive compared with the standard method. In Wisconsin, it is estimated that the additional time required to sample QNS specimens instead of disregarding them is approximately 1 hour per 100 QNS specimens (Gary L. Hoffman, unpublished data). In other states, the additional time required to punch specimens using the alternate method will vary, depending on the number of QNS specimens and the methods used to punch specimens.

Since 1988, HIV-1 seroprevalence surveys that use dried neonatal blood specimens have become a major component of the national effort to track the epidemic of HIV-1 infection. These important studies are being relied on to reveal trends in HIV-1 infection among childbearing women and to estimate the incidence of vertically transmitted HIV-1 infection in the United States. Information obtained from these surveys can be used in planning activities to evaluate the need for and effectiveness of specific HIV-1-related prevention programs and to anticipate future demand for HIV-1-related medical and social services. Because the decisions derived from these activities are important, every effort must be made to understand and eliminate bias that may affect the accuracy of HIV-1 seroprevalence estimates. The QNS effect is but one of several potential biases inherent in these surveys.[16]

Using multiple 1/8-in blood spots to sample dried neonatal blood specimens that are of insufficient quantity to analyze using standard 1/4-in blood spots can be an effective way of ensuring the accuracy of HIV-1 seroprevalence estimates. Thus, public health agencies and laboratories involved with surveys to estimate HIV-1 seroprevalence among childbearing women in other states and regions may wish to incorporate the use of multiple 1/8-in blood spots into their survey protocols. □

References

1. *HIV/AIDS Surveillance Report.* Atlanta, Ga: Centers for Disease Control; January 1991:1–15.
2. Ellerbrock TV, Bush TJ, Chamberland ME, Oxtoby MJ. Epidemiology of women with AIDS in the United States, 1981 through 1990. *JAMA.* 1991;265:2971–2975.
3. Pappaioanou M, George JR, Hannon WH, et al. HIV seroprevalence surveys of childbearing women: objectives, methods, and uses of data. *Public Health Rep.* 1990;105: 147–152.
4. Gwinn M, Pappaioanou M, George JR, et al. Prevalence of HIV infection in childbearing women in the United States: surveillance using newborn blood samples. *JAMA.* 1991;265:1704–1708.
5. Hoff R, Berardi VP, Weiblen BJ, Mahoney-Trout L, Mitchel ML, Grady GF. Seroprevalence of human immunodeficiency virus among childbearing women: estimation by testing samples of dried blood from newborns. *N Engl J Med.* 1988; 318:525–530.
6. Novick LF, Berns D, Stricof R, Stevens R, Pass K, Wethers J. HIV seroprevalence in newborns in New York State. *JAMA.* 1989;261:1745–1750.
7. Centers for Disease Control. Interpretation and use of Western blot assay for serodiagnosis of human immunodeficiency virus type 1 infections. *MMWR.* 1989;38(suppl S–7):1–7.
8. Breslow NE, Day NE. *Statistical Methods in Cancer Research.* Lyon, France: International Agency for Cancer Research; 1980. IARC scientific publication 32.
9. *Maternal and Child Health Statistics.* Madison, Wis: Wisconsin Department of Health and Social Services; 1989:34.
10. Morse DL, Lessner L, Medvesky MG, Glebatis DM, Novick LF. Geographic distribution of newborn HIV seroprevalence in relation to four sociodemographic variables. *Am J Public Health.* 1991;81(suppl):25–29.
11. Fanaroff AA, Martin RJ, eds. *Neonatal and Perinatal Medicine.* 4th ed. St. Louis, Mo: CV Mosby Company; 1987.
12. Neerhof MG, MacGregor SN, Retsky SS, Sullivan TP. Cocaine abuse during pregnancy: peripartum prevalence and perinatal outcome. *Am J Obstet Gynecol.* 1989; 161:633–638.
13. McCalla SA, Minkoff HL, Feldman J, et al. The biologic and social consequences of perinatal cocaine use in an inner-city population: results of an anonymous cross-sectional study. *Am J Obstet Gynecol.* 1991;164:625–630.
14. Rogers M. Perinatal infection. In: Kaslow RA, Francis DP, eds. *Epidemiology of AIDS.* New York, NY: Oxford University Press; 1989:231–241.
15. Oxtoby MJ. Perinatally acquired HIV infection. In: Pizzo PA, Wilfert CM, eds. *Pediatric AIDS: The Challenge of HIV Infection in Infants, Children and Adolescents.* Baltimore, Md: Williams and Wilkins; 1991:3–21.
16. Pass KA, Schedlbauer LM, MacCubbin PA, Glebatis DM. Comparison of newborn screening records and birth certificates to estimate bias in newborn HIV serosurveys. *Am J Public Health.* 1991; 81(suppl):22–24.

The Disclosure of Celebrity HIV Infection: Its Effects on Public Attitudes

Seth C. Kalichman, PhD, and Tricia L. Hunter

ABSTRACT

Objectives. **Despite the magnitude of the acquired immunodeficiency syndrome (AIDS) epidemic, studies have shown low levels of public concern about human immunodeficiency virus (HIV)/AIDS. We investigated the effects of celebrity disclosure of HIV infection on the AIDS-related perceptions of urban men.**

Methods. **Measures of AIDS-related perceptions were collected from 361 men waiting for mass transportation in downtown Chicago; 252 were assessed at three time points prior to and 109 were assessed at two time points after professional basketball star Earvin "Magic" Johnson's disclosure of his HIV infection.**

Results. **Significant increases in concern about AIDS, interest in AIDS information, and talking with friends about AIDS occurred after celebrity disclosure of HIV infection.**

Conclusions. **Celebrity disclosure of HIV seropositivity demonstrated a marked change in AIDS-related perceptions among the men surveyed. Changes in awareness due to celebrity disclosure may lead to increased readiness to reduce risk and could be viewed as a window of opportunity for HIV prevention efforts. (*Am J Public Health.* 1992; 82:1374–1376)**

Introduction

More than 135 000 persons in the United States have died from complications of acquired immunodeficiency syndrome (AIDS), 209 000 have been diagnosed with AIDS,[1] and more than 1 million are believed to be infected with the human immunodeficiency virus (HIV).[2] Heterosexually contracted HIV has increased by more than 8% from 1991 to 1992.[1,3] However, public perceptions about HIV/AIDS have not paralleled the magnitude of the epidemic. Surveys have shown a lack of concern about AIDS among the general public.[4–8] Consequently, the social climate in the United States has not facilitated population-wide HIV-risk reduction.

One potential agent of change in the social climate regarding AIDS is disclosure of celebrity HIV seropositivity. If we assume that identification with a target celebrity results in personal relevance, the disclosure of a celebrity's HIV seropositivity should affect one's self-perceived susceptibility. A recent occurrence of just such an event afforded us the opportunity to investigate this hypothesis. In a national press conference on November 7, 1991, basketball star Earvin "Magic" Johnson disclosed that he was HIV seropositive.[9] The present study investigated the effects of Johnson's self-disclosure on the AIDS-related perceptions of urban men.

Methods

Study Design

To conduct a community assessment of men's perceptions of HIV/AIDS, we collected data at main arteries of the mass transit system in Chicago, which ranks eighth in reported AIDS cases among U.S. metropolitan areas.[1] Data were collected in the summer and fall of 1991 (n = 169), and 10 and 5 days before Magic Johnson's disclosure (on October 28 [n = 58] and November 2 [n = 25], respectively), providing three baseline data points; data were also collected 3 and 10 days after Johnson's disclosure (on November 10 [n = 49] and November 16 [n = 60], respectively), with questions added to assess specific knowledge about his self-disclosure.

Measures and Procedure

Measures were derived from questions related to risk perceptions, interest in AIDS information (see Table 1), frequency of talking with friends about AIDS, and risk behavior histories. All questions were reviewed for content and clarity by an HIV health educator and were pilot tested prior to the study.

Men were approached as they waited for mass transportation. Asked to complete surveys about AIDS, potential subjects were told that the survey would take approximately 10 minutes, could be completed while they waited for their train, and contained personal and sensitive questions. Those agreeing to participate were provided with both a survey attached to a clipboard and a pencil. After returning the survey, participants were given a flyer that described the study, pro-

Seth C. Kalichman is with the Medical College of Wisconsin in Milwaukee, and Tricia L. Hunter is with Loyola University of Chicago, Ill.

Requests for reprints should be sent to Dr. Seth C. Kalichman, PhD, Department of Psychiatry and Mental Health Sciences, Community Health Behavior Program, Medical College of Wisconsin, 8701 Watertown Plank Rd, Milwaukee, WI 53226.

This paper was submitted to the Journal November 27, 1991, and accepted with revisions May 22, 1992.

vided accurate information about HIV/AIDS transmission, and included a phone number for further information. Participants were also compensated with a city mass transit token.

Sample Description

Of the 435 men who were approached, 361 (83%) participated. The respondents were 28.8 years of age (SD = 8.0); 94% completed high school; 42% reported an annual income of less than $10 000; 55% were African American; 84% were unmarried; 51% had two or more sexual partners in the past year; 7% had histories of injection drug use; 12% had sexual partners who used injection drugs; 23% had had sex with a prostitute; 19% had had homosexual contact; 32% had a history of a sexually transmitted disease; 33% had been tested for HIV antibodies; and 7% were HIV seropositive. Nonparametric and parametric statistical tests did not show any differences in sample characteristics across the five assessment points. Although it cannot be assumed that the participants are representative of men living in Chicago, their demographic characteristics approximate those of Chicago urban communities with high AIDS incidence rates.[10,11]

Results

Analyses of variance on the five assessment times with respect to AIDS-related perceptions indicated significant differences in concern over an acquaintance getting AIDS (F [4, 353] = 3.23, $P < .01$); the assessment taken 3 days after Johnson's disclosure demonstrated greater levels of concern than all three baseline points, and the assessment taken 10 days after showed greater concern than two of the baseline points. Self-concern and self-estimated risk of HIV infection also showed a similar pattern of increased concern but were not statistically significant (see Table 1).

Interest in getting more AIDS information changed significantly from baseline to postdisclosure (F [4, 341] = 2.96, $P < .02$). Of men assessed 3 and 10 days after disclosure, 58% and 56%, respectively, were very interested in getting more AIDS information, as compared with 42%, 38%, and 32% of men at the three successive baseline assessments (see Table 1). Similarly, men were more likely to discuss AIDS with friends after the disclosure (χ^2 [4] = 10.89, $P < .03$) (see Figure 1).

TABLE 1—Mean (SD) for Perceived Susceptibility and Interest Variables across Assessment Times before and after Magic Johnson's Self-Disclosure

	Before Disclosure			After Disclosure		95% Confidence Interval
	6 Months	10 Days	5 Days	3 Days	10 Days	
I am concerned that I will get AIDS[a]	2.36 (1.13)	2.38 (1.02)	2.48 (.96)	2.63 (1.11)	2.65 (1.04)	2.35, 2.57
I am concerned that someone I know will get AIDS[a]	2.84 (1.07)	2.77 (.99)	2.68 (.95)	3.27 (.94)	3.15 (.94)	2.83, 3.04
What is your risk for getting the AIDS virus?[b]	2.35 (1.11)	2.24 (.99)	2.28 (.98)	2.35 (1.07)	2.78 (1.24)	2.28, 2.52
How interested are you in getting more information about AIDS?[c]	3.21 (.84)	3.13 (.84)	3.04 (.89)	3.52 (.65)	3.44 (.77)	3.19, 3.36

[a]Ratings on a 4-point scale: 1 = not true for me; 4 = very true for me.
[b]Ratings on a 5-point scale: 1 = no risk at all; 5 = extremely high risk.
[c]Ratings on a 4-point scale: 1 = not at all interested; 4 = very interested.

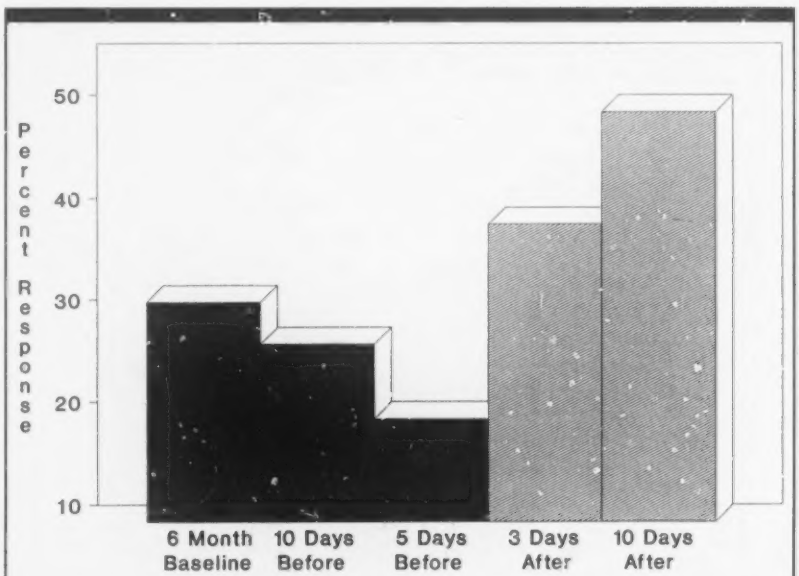


FIGURE 1—Percentage indicating that they "often" talked with friends about AIDS, across five assessment points.

All men surveyed at both postdisclosure assessments indicated that they had heard about Magic Johnson. While 86% had discussed the disclosure 3 days after, 100% had talked about it by 10 days after. Comparisons between African-American and Caucasian men showed no differences in talking with friends about the disclosure. However, significant differences were found on the question "Are you more concerned about AIDS since hearing about Magic Johnson?" (Measurement was taken using a 4-point scale ranging from "Not at all" to "Much more concerned.") African-American men indicated that Johnson's disclosure increased their concern about AIDS to a greater extent than did Caucasian men both 3 days after (t[30] = 4.25, $P < .01$) and 10 days after (t[35] = 3.43, $P < .01$) (see Figure 2).

Discussion

The shift in social climate identified in the present study appears to be a function of Magic Johnson's self-disclosure, although lack of experimental control does

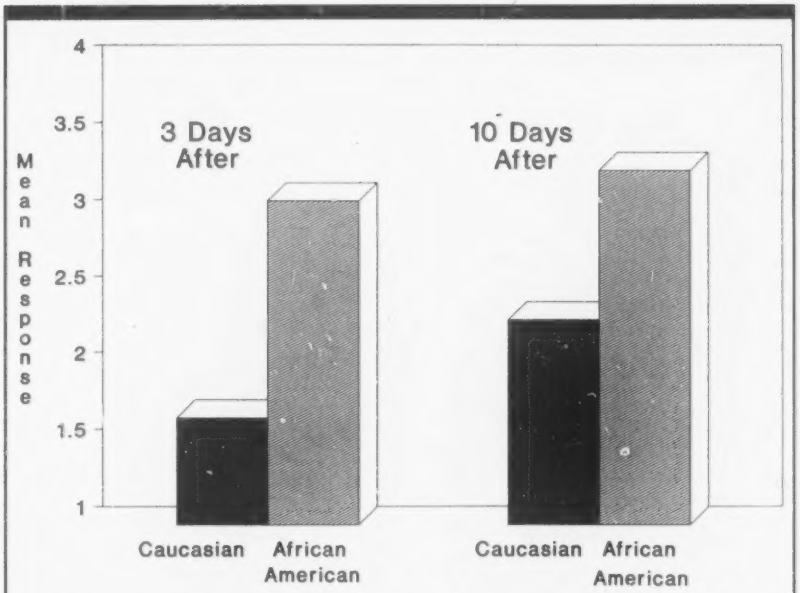


FIGURE 2—Mean responses for Caucasian and African-American men reflecting the impact of Magic Johnson's disclosure on their concern about AIDS.

not allow for the exclusion of alternative causal factors. Substantially more African-American men than Caucasian men indicated that the disclosure influenced their concern about AIDS. Because mental representations of persons afflicted with a health problem are related to their perceptions of susceptibility,[12] the impact of celebrity disclosures may vary across subpopulations, depending on group associations with the target celebrity.

The impact of celebrity disclosures on urban men's perceptions sets the stage for risk-reducing behaviors. We conclude that Magic Johnson's self-disclosure caused some men to shift from precontemplation toward contemplation of self-change.[13,14] With continued awareness and information, men may engage in self-reevaluation processes that ultimately inspire behavioral changes.[13] Thus, disclosure of celebrity HIV seropositivity provides a natural enhancement for AIDS prevention messages. Given the likely, albeit unpredictable, occurrence of such disclosures in the future, prevention programs would do well to capitalize on their immediate impact. For example, concentrated education and prevention campaigns may be prepared in advance and held until a celebrity disclosure occurs. Additionally, further research is needed to understand better the link between celebrity disclosure and risk perceptions, as well as the relationship between any resultant change in the social climate and changes in risk behavior. Finally, social risks should be minimized to encourage disclosure, and celebrity messages should be tailored to maximize their social impact. □

Acknowledgments

This study was funded with a Research Support Grant from Loyola University of Chicago. The authors appreciate the assistance of Moira Kalichman, Delores Heard, the Chicago Department of Health, and the men of Chicago who participated. Jeffrey A. Kelly and Robert L. Russell are thanked for providing helpful comments.

References

1. *HIV/AIDS Surveillance.* Atlanta, Ga: Centers for Disease Control; February 1992.
2. Centers for Disease Control. HIV prevalence estimates and AIDS case projections for the United States: report based on a workshop. *MMWR.* 1990;39:30.
3. Holmes K, Karon J, Kreiss J. The increasing frequency of heterosexually acquired AIDS in the United States, 1983–88. *Am J Public Health.* 1990;80:858–863.
4. Aruffo J, Coverdale J, Vallbona C. AIDS knowledge in low-income and minority populations. *Public Health Rep.* 1991;106: 115–119.
5. DiClemente R, Boyer C, Morales E. Minorities and AIDS: knowledge, attitudes, and misconceptions among Black and Latino adolescents. *Am J Public Health.* 1988;78:55–57.
6. Kappel S, Vogt R, Brozicevic M, Kutzko D. AIDS knowledge and attitudes among adults in Vermont. *Public Health Rep.* 1989;104:388–391.
7. Kalichman SC, Hunter T, Kelly J. Perceptions of AIDS susceptibility among minority and nonminority women at risk for HIV infection. *J Consult Clin Psychol.* In press.
8. Gerbert B, Sumser J, Maguire B. The impact of who you know and where you live on opinions about AIDS and health care. *Soc Sci Med.* 1991;32:677–681.
9. Johnson E, Johnson R. I'll deal with it (interview with Earvin Magic Johnson). *Sports Illustrated.* November 18, 1991:16–45.
10. *AIDS Chicago.* Chicago, Ill: Chicago Department of Health, Surveillance Program, Office of AIDS Prevention; June 1991.
11. Chicago Fact Book Consortium. *Local Community Fact Book Chicago Metropolitan Area.* Chicago, Ill: Chicago Review Press; 1984.
12. Weinstein ND. Perceptions of personal susceptibility to harm. In: Mays V, Albee G, Schneider S, eds. *Primary Prevention of AIDS: Psychological Approaches.* Newbury Park, Calif: Sage; 1989.
13. Prochaska JO, DiClemente C. Stages and processes of self-change of smoking: toward an integrative model of change. *J Consult Clin Psychol.* 1983;51: 390–395.
14. Prochaska JO, Velicer W, DiClemente C, Fava J. Measuring processes of change: applications to the cessation of smoking. *J Consult Clin Psychol.* 1988;56:520–528.

Public Health Briefs

Legal Blindness among 10-Year-Old Children in Metropolitan Atlanta: Prevalence, 1985 to 1987

Carolyn D. Drews, PhD, Marshalyn Yeargin-Allsopp, MD, Catherine C. Murphy, MPH, and Pierre Decoufle, ScD

ABSTRACT

The prevalence of legal blindness in 10-year-olds in metropolitan Atlanta was 6.8 per 10 000 during 1985 to 1987. The prevalence was 8.8 per 10 000 in Black boys, 8.6 per 10 000 in White boys, 6.7 per 10 000 in White girls, and 1.8 per 10 000 in Black girls. Retinopathy of prematurity was the most common known cause (1.0 per 10 000). Of the 61 cases, 40 had other disabilities, including 14 with mental retardation, cerebral palsy, and epilepsy. The low prevalence among Black girls and the frequent occurrence of blindness with other disabilities are noteworthy. (*Am J Public Health.* 1992;82:1377–1379)

Introduction

There is a lack of current data on the prevalence of childhood blindness in the United States. This paper examines three aspects of childhood blindness in Atlanta during 1985 to 1987: its prevalence, causes, and association with other developmental disabilities.

Materials and Methods

The Metropolitan Atlanta Developmental Disabilities Study estimated the prevalence of five developmental disabilities in 10-year-olds living in five metropolitan Atlanta counties during 1985 to 1987.[1]

Blind children were identified by reviewing records at public schools, hospitals, health and mental health service agencies, and public and private social service agencies. Blindness was defined as a best corrected visual acuity of 20/200 or worse in the better eye[2] or a statement by a trained person (i.e., ophthalmologist or optometrist) that the child was blind, since measuring visual acuity in young children, especially those with other disabilities, is often difficult. Age at documentation is defined as the earliest age at which records available to us stated that the child was blind or had a best corrected visual acuity of 20/200 or worse.

Etiology was defined as the underlying condition that led to blindness. Etiology was considered to be unknown if an underlying condition could not be determined from the available records even if the child was at high risk for blindness (e.g., due to prematurity). Age at onset was defined as the age at the time of the presumed etiologic event.

Prevalence rates and their 95% confidence intervals (CI) are reported per 10 000 10-year-olds.[3] Using data provided by the Georgia Office of Planning and Budget, we estimated that 89 534 ten-year-olds lived in Atlanta during 1985 to 1987. Race was defined as White or Black since all cases were either White or Black and 96% of children in Atlanta who are not White are Black.

Results

We identified 61 blind children (6.8 per 10 000) (Table 1). The prevalence of blindness was slightly higher in Whites than in Blacks but was higher in boys than in girls. However, the magnitude of the gender difference varied substantially by race.

Retinopathy of prematurity was the most common known cause of blindness (1.0 per 10 000; 95% CI = 0.5, 1.9) (Table 2). Thirteen children had an eye malformation. Two other children had cataracts resulting from prenatal infections. Thus, one quarter (15) of the cases were known

Carolyn D. Drews is with the Division of Epidemiology of the Emory University School of Public Health and the Department of Ophthalmology of the Emory University School of Medicine, Atlanta, Ga. Marshalyn Yeargin-Allsopp and Pierre Decoufle are with the Division of Birth Defects and Developmental Disabilities, National Center for Environmental Health and Injury Control, Atlanta, Ga. Catherine C. Murphy is with the Office of Epidemiology, Georgia Department of Human Resources, Atlanta, Ga.

Requests for reprints should be sent to Carolyn D. Drews, PhD, Emory University School of Public Health, 1599 Clifton Rd, NE, Atlanta, GA 30329.

This paper was submitted to the Journal December 3, 1991, and accepted with revisions May 19, 1992.

TABLE 1—Prevalence of Legal Blindness among 10-Year-Old Children (per 10 000) by Race and Sex, Metropolitan Atlanta, 1985 to 1987

	White			Black			Total		
Sex	n	Rate	95% CI	n	Rate	95% CI	n	Rate	95% CI
Male	25	8.6	5.6, 12.7	15	8.8	4.9, 14.5	40	8.7	6.2, 11.8
Female	18	6.7	4.0, 10.5	3	1.8	0.4, 5.3	21	4.9	3.0, 7.5
Total	43	7.7	5.5, 10.3	18	5.4	3.2, 8.5	61	6.8	5.2, 8.8

TABLE 2—Etiology of Blindness among 61 Legally Blind 10-Year-Old Children, Metropolitan Atlanta, 1985 to 1987

	No.	%[a]
Congenital eye malformations	13	21.3
Genetic or hereditary condition	4	6.6
Albinism	1	1.6
Retinitis pigmentosa	1	1.6
Tuberous sclerosis[b]	1	1.6
Stargardt's syndrome	1	1.6
Infection	5	8.2
Prenatal	2	3.3
Postnatal	3	4.9
Retinopathy of prematurity	9	14.2
Other		
Anoxia/hypoxia	2	3.3
Central nervous system defect	4	6.6
Cerebrovascular accident	1	1.6
Head trauma	4	6.6
High myopia	1	1.6
Hydrocephaly	1	1.6
Hypoglycemia[c]	1	1.6
Neoplasm	2	3.3
Optic nerve atrophy	3	4.9
Radiation[d]	1	1.6
Stevens-Johnson syndrome[e]	1	1.6
Mixed etiology[f]	1	1.6
Missing[g]	8	13.1

[a]Does not add to 100% because of rounding.
[b]Tuberous sclerosis caused hydrocephaly and optic atrophy.
[c]Hypoglycemia of unknown etiology caused coma with subsequent cortical blindness.
[d]Radiation optic neuropathy was secondary to treatment for nasopharyngeal rhabdomyosarcoma.
[e]Stevens-Johnson syndrome caused corneal scarring.
[f]Blindness due to congenital glaucoma in one eye and myopia in the other.
[g]No information available on the etiology of blindness.

to have an eye malformation. Information on etiology was missing for eight cases.

Two thirds (39) of the cases had the condition that blinded them by the time they were 1 month old, but few were documented to be blind in infancy (Table 3).

Two thirds of the cases had at least one of four other disabilities; 40% (25) had at least two (Table 4). Fourteen (23%) children had mental retardation, cerebral palsy, and epilepsy (1.6 per 10 000; 95% CI = 0.9, 2.6).

Discussion

The prevalence of blindness in 10-year-olds in Atlanta is slightly higher than previously reported rates for developed countries (3.4 to 6.2 per 10 000).[4-8] Variations in the reported rates may be due to instability of rates based on small numbers or may represent real differences in occurrence, especially if survival of high-risk infants differs. Differences in ascertainment methods may also create apparent differences in prevalence rates. Cases may be underascertained if identified solely through schools for the blind or through blindness registries since these programs are usually voluntary.[4] Only 6 of 61 cases in this study had attended schools for the blind; 21 of the children with multiple disabilities had attended special education programs that were not designed specifically for blind children.

As in other studies, boys were more likely to be blind than girls.[5,6,8,9] However, six of the nine children with retinopathy of prematurity were girls. This finding could be due to small numbers, better survival of low-birthweight girls than boys,[10] or gender-based differences in birthweight.[6]

The overall prevalence of blindness was slightly higher in Whites than in Blacks, but the pattern differed by gender. Reasons for this finding are not clear, and the rates are based on small numbers. Few previous studies have reported race-specific or race-and-sex-specific rates of childhood blindness. An early study found that the prevalence was 20% lower among Whites than among other races, but the rates were not presented by gender.[4]

The low prevalence of blindness among Black girls is striking and should be confirmed in other populations. Ascertainment could differ by race or socioeconomic status but should not vary by gender. Thus, we do not believe that we have undercounted cases among Black girls. The prevalence might be low in Black girls if they are less likely to have conditions that put them at risk of blindness or if those at high risk are less likely to survive than other children.

As in previous studies, about two thirds of our blind children had other disabilities.[7,9,11] Blindness, mental retardation, epilepsy, and cerebral palsy occurred with an unexpectedly high frequency in this population. The occurrence of this combination of disabilities should be examined in other populations.

In earlier studies, 7% to 14% of blind children were hearing impaired, with most cases attributed to rubella.[8,9,11] In contrast, only three (5%) of our cases were hearing impaired; one case was attributed to congenital rubella syndrome. Thus, rubella appears to be a less important cause of blindness among children born in the mid-1970s than in previous birth cohorts.[12]

Our results underscore the concern that improved survival of low-birthweight infants may increase the prevalence of

TABLE 3—Age at Onset of Etiology and Age at Earliest Documentation of Blindness for 61 Legally Blind 10-Year-Old Children, Metropolitan Atlanta, 1985 to 1987

	Age at Onset, No. (%)	Age at Documentation, No. (%)
Age known	53 ...	61 ...
<1 month	39 (63.9)	1 (1.6)
1–11 months	6 (9.8)	8 (13.1)
1–4 years	4 (6.6)	21 (34.4)
5–10 years	4 (6.6)	31 (50.8)
Age unknown	8[a] (13.1)	0 ...

[a]One of these children was known to be blind by 10 months of age.

TABLE 4—Number of Legally Blind 10-Year-Old Children with Selected Other Disabilities, Metropolitan Atlanta, 1985 to 1987

Other Disabilities	No.	%
None	21	34.4
Mental retardation	9	14.8
Epilepsy	3	4.9
Cerebral palsy	2	3.3
Hearing impairment	1	1.6
Mental retardation and epilepsy	4	6.6
Mental retardation and cerebral palsy	3	4.9
Mental retardation and hearing impairment	2	3.3
Epilepsy and cerebral palsy	2	3.3
Epilepsy and hearing impairment	0	...
Cerebral palsy and hearing impairment	0	...
Mental retardation, epilepsy, and cerebral palsy	14	23.0
Mental retardation, cerebral palsy, and hearing impairment	0	...
Mental retardation, epilepsy, and hearing impairment	0	...
Epilepsy, cerebral palsy, and hearing impairment	0	...
Mental retardation, epilepsy, cerebral palsy, and hearing impairment	0	...

blindness due to retinopathy of prematurity since this condition accounted for 15% of our cases.[13–17]

Most of the conditions that caused blindness originated prenatally, and one quarter of the blind children had a congenital eye defect. These proportions might have been higher if our data on etiology had been more complete. However, most cases of blindness were first documented after age 4. The lag between age at onset and documentation may be related to our methodology or difficulties in measuring vision in very young children, and does not necessarily indicate a delay in treatment of childhood visual impairments. These findings emphasize the importance of early intervention and of identifying ways to improve the visual outcome of children born with conditions that put them at risk of being blind. □

Acknowledgments

This report was supported by funds from the Comprehensive Environmental Response, Compensation, and Liability Act trust fund through an interagency agreement with the Agency for Toxic Substances and Disease Registry, Public Health Service, US Department of Health and Human Services. Dr. Drews is supported in part by P30 EY0630 (a National Institutes of Health departmental core grant).

We especially thank Drs. Muin Khoury, Michelle Kiely, and Robert Sperduto for reviewing an early version of this manuscript.

References

1. Yeargin-Allsopp M, Murphy CC, Oakley GP, et al. A multiple-source method for studying the prevalence of developmental disabilities in children: the Metropolitan Atlanta Developmental Disabilities Study. *Pediatrics.* 1992;89:624–630.
2. Vaughan D, Asbury T. *General Ophthalmology.* 11th ed. Norwalk, Conn: Appleton-Century-Crofts; 1986:393.
3. Person ES, Hartley HO, eds. *Biometrika Tables for Statisticans.* 3rd ed. Cambridge, England: Cambridge University Press; 1970;1.
4. Goldstein H. The reported demography and causes of blindness throughout the world. *Adv Ophthalmol.* 1980;40:1–99.
5. National Society to Prevent Blindness. *Vision Problems in the United States. Data Analysis: Definitions, Data Sources, Detailed Data Tables, Analysis, and Interpretation.* New York, NY: National Society to Prevent Blindness; 1980.
6. Hatfield EM. Why are they blind? *Sight Saving Rev.* 1975;45(1):3–22.
7. Stewart-Brown SL, Haslum MN. Partial sight and blindness in children of the 1970 birth cohort at 10 years of age. *J Epidemiol Community Health.* 1988;42:17–23.
8. Halldorsson S, Bjornsson G. Childhood blindness in Iceland: a study of legally blind and partially seeing children in Iceland, 1978. *Acta Ophthalmol.* 1980;58: 237–242.
9. Robinson GC, Jan JE, Kinnis C. Congenital ocular blindness in children, 1945 to 1984. *AJDC.* 1987;141:1321–1324.
10. Hogue CH, Strauss LT, Buehler JW, Smith JC. Overview of the National Infant Mortality Surveillance Project, 1980. *MMWR.* 1989;38:1–46.
11. Rosenberg T. Visual impairment in Danish Children, 1985. *Acta Ophthalmol.* 1987;65: 110–117.
12. Cochi SL, Edmonds LE, Dyer K, et al. Congenital rubella syndrome in the United States, 1970–1985: on the verge of elimination. *Am J Epidemiol.* 1989;29: 349–361.
13. Stark DJ. Retinopathy of prematurity: a second epidemic? *Med J Aust.* 1989;150: 289–290.
14. Gibson DL, Sheps SB, Schechter MT, Wiggins S, McCormick AQ. Retinopathy of prematurity: a new epidemic? *Pediatrics.* 1989;83:486–492.
15. Gibson DL, Sheps SB, Uh SH, Schechter MR, McCormick AQ. Retinopathy of prematurity-induced blindness: birth weight-specific survival and the new epidemic. *Pediatrics.* 1990;86:405–412.
16. Valentine PH, Jackson JC, Kalina RE, Woodrum DE. Increased survival of low birth weight infants: impact on the incidence of retinopathy of prematurity. *Pediatrics.* 1989;84:442–445.
17. Keith CG, Smith ST, Lansdell BJ. Retrolental fibroplasia: a study of the incidence and etiological factors, 1977–1979. *Med J Aust.* 1981;2:589–592.

Sudden Infant Death Syndrome and Maternal Smoking

Michael H. Malloy, MD, MS, Howard J. Hoffman, MA, and Donald R. Peterson, MD, MPH

ABSTRACT

Data from Missouri for the period 1980 to 1985 suggest a dose–response relationship between smoking during pregnancy and the incidence of sudden infant death syndrome (SIDS). However, data from the National Institute of Child Health and Human Development SIDS Cooperative Epidemiological Study did not support a dose–response relationship. Neither the Missouri data nor the Cooperative Study data support a relationship between the age of occurrence of SIDS and smoking during pregnancy. (*Am J Public Health.* 1992;82:1380–1382)

Introduction

A number of socioeconomic and demographic factors are associated with an increased risk of sudden infant death syndrome (SIDS), but few readily amenable exposure factors have been identified.[1] Reports of an association between maternal cigarette smoking during pregnancy and the subsequent occurrence of SIDS, however, do offer the possibility that reducing maternal smoking could result in a decline in the SIDS rate.[2–4] Whether the association between SIDS and a history of maternal smoking during pregnancy is biological in nature or a proxy for maternal behavior is not clear. Haglund and Cnattingius reported that infants born to women who smoke during pregnancy die earlier with SIDS than infants of nonsmokers and that a dose–response association exists between maternal smoking and the risk of SIDS.[5] Their report supports the plausibility of a biological mechanism. We attempted to confirm their observations in two large US data sets.

Methods

The Missouri data set is a linked birth and death certificate file for the period 1980 to 1985. There were 684 SIDS deaths in Missouri during this period. Complete data were available on 636 of these cases and 425 326 live births, an overall rate of 1.6/1000 live births (1.3 non-Black vs 3.1 Black). For this analysis, we excluded infants of gestational age less than 24 weeks or greater than 45 weeks and excluded cases of SIDS in which the age of death was less than 7 days. Seventy-six percent of the SIDS deaths reported between 1980 and 1985 were autopsied, but no other special review process was carried out to confirm the validity of the assignment of the diagnosis of SIDS. Additional information obtained from the birth certificate included race, maternal age, parity, maternal education, and marital status. Maternal smoking information was obtained from the birth certificate and was reported in terms of packs of cigarettes smoked per day.

The National Institute of Child Health and Human Development SIDS Cooperative Epidemiological Study was a case-control study conducted between October 1978 and December 1979 that accrued cases of SIDS across six geographic regions.[6] The SIDS rate for this study was 1.75/1000 live births (1.10 non-Black vs 3.65 Black). All cases of SIDS were autopsied, and a panel of three expert pathologists reviewed each of the cases to determine the reliability of the diagnosis. Living controls (Control A) were selected from regional birth certificate files to match regional cases on the age of death. A second living control (Control B) was matched to the SIDS case within each region on age, race, and birthweight (≤2500 g). There were 757 SIDS cases and 1514 controls available for this analysis. Mothers of SIDS cases were interviewed 2 to 5 weeks after the child died. Mothers of cases and controls were asked during the interview how many cigarettes per day they smoked during the first, second, and third trimesters. For this analysis, information on smoking in the first trimester was used.

Analyses were carried out to determine whether the SIDS infants' age of death was less when the mother smoked than when the mother did not smoke by using general chi-square tests of association. The relationship between the quantity smoked and infant mortality was analyzed by logistic regression.[7] Two-way interactions between smoking and covariates in the logistic models were exam-

Michael H. Malloy and Howard J. Hoffman are with the Division of Epidemiology, Statistics and Prevention Research, National Institute of Child Health and Human Development, Bethesda, Md. Michael H. Malloy is also with the Department of Pediatrics, University of Texas Medical Branch, Galveston. Donald R. Peterson is with the Department of Biostatistics, University of Washington, Seattle.

Requests for reprints should be sent to Michael H. Malloy, MD, MS, Department of Pediatrics, Rte E-26, University of Texas Medical Branch, Galveston, TX 77550.

This paper was submitted to the Journal October 8, 1991, and accepted with revisions May 26, 1992.

ined, but none were significant. Analyses were done with SAS.[8]

Results

In the Missouri population and in the Cooperative Study, the distribution of the age of SIDS death for smokers and nonsmokers appears quite similar (Figures 1 and 2). There were no statistically significant associations by general chi-square analysis. The median age of death for both nonsmokers and smokers was 11 weeks. This contrasts with the observation from Sweden of a median age of SIDS death among nonsmokers of 10.4 weeks compared with 8 weeks among smokers.[5]

The dose–response relationship between smoking and SIDS was examined by logistic regression (Table 1). The odds ratios were adjusted by logistic regression for race, parity, age of the mother, maternal education, marital status, and sex of the infant. For Missouri, the adjusted odds ratios were 1.98 for light smokers compared with nonsmokers and 2.86 for heavy smokers compared with nonsmokers. The odds ratio for heavy smokers compared with light smokers was 1.41 with a 95% confidence interval (CI) that did exclude 1.0.

For the Cooperative Study, the adjusted odds ratios, with the reference group being nonsmokers and using Control A (the controls matched on age alone), were 2.49, 3.24, and 3.05 for smokers of less than 10, 10 through 19, and 20 or more cigarettes per day, respectively. For Control B (the controls matched on age, birthweight, and race), the odds ratios were 2.68, 2.65, and 2.60, respectively.

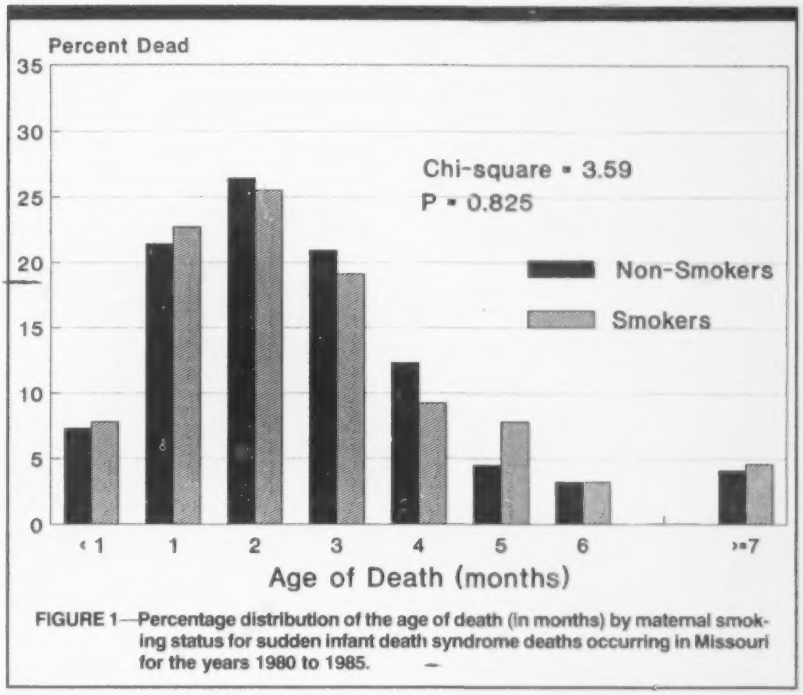


FIGURE 1—Percentage distribution of the age of death (in months) by maternal smoking status for sudden infant death syndrome deaths occurring in Missouri for the years 1980 to 1985.

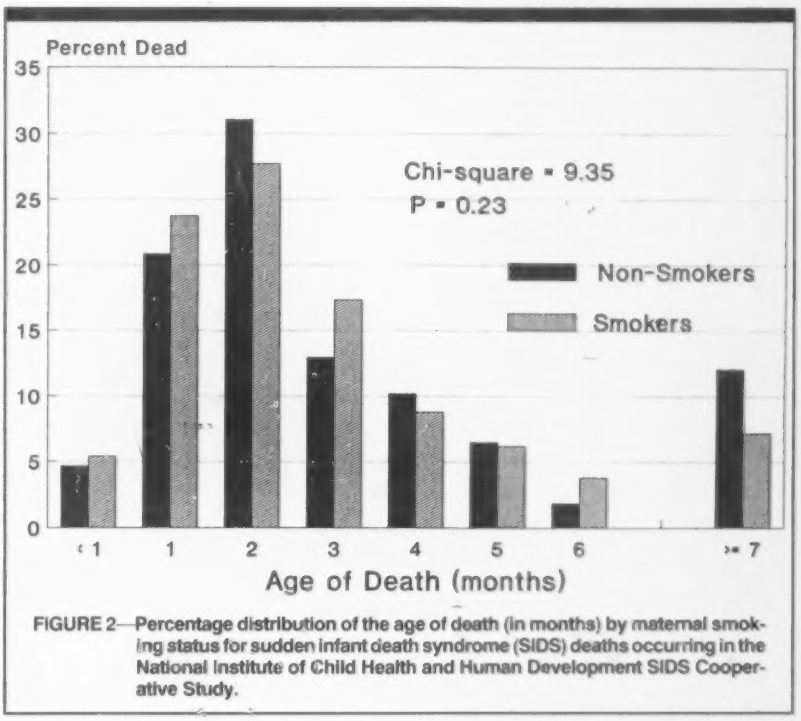


FIGURE 2—Percentage distribution of the age of death (in months) by maternal smoking status for sudden infant death syndrome (SIDS) deaths occurring in the National Institute of Child Health and Human Development SIDS Cooperative Study.

Discussion

We were unable to confirm the Swedish observation of a relationship between the age of death and a history of maternal smoking during pregnancy. In the Missouri population, we did observe a relationship between the quantity smoked and the risk for SIDS. Data from the Cooperative Study, however, failed to support a dose–response relationship. Differences in results between the two populations may be due to recall bias. In the Cooperative Study women reported smoking after the death of the infant, while in the Missouri population smoking information was reported on the birth certificate prior to the death of the infant. However, the direction of this potential bias is uncertain.

Confirmation of the Swedish report of an earlier age of death among infants of smokers would support the hypothesis that the relationship between maternal smoking and SIDS is biologic. However, we were unable to demonstrate this relationship with two large US data sets. Nevertheless, further research into biologic mechanisms and life-style factors by which maternal smoking might increase the risk of SIDS should be pursued.

TABLE 1—Adjusted Odds Ratios[a] for Sudden Infant Death Syndrome (SIDS) from the Missouri and NICHD[b] SIDS Cooperative Epidemiological Study Populations

Population/No. of Cigarettes Smoked Daily	Adjusted Odds Ratio	95% CI
Missouri		
Nonsmoker	1.00	...
<1 pack	1.98	1.64, 2.41
≥1 pack	2.86	2.32, 3.50
≥1 pack vs <1 pack	1.41	1.13, 1.76
Cooperative study		
Control A		
Nonsmoker	1.00	...
<10	2.49	1.79, 3.46
10–19	3.24	2.31, 4.55
≥20	3.05	2.24, 4.15
Control B		
Nonsmoker	1.00	...
<10	2.68	1.94, 3.71
10–19	2.65	1.91, 3.67
≥20	2.60	1.93, 3.52

[a]Odds ratios adjusted for maternal age (< 20 vs ≥20 years), race (Black vs non-Black), parity (≥2 vs <2), educational status (<12 vs ≥12 years), marital status (single vs married), and sex of the infant.
[b]National Institute of Child Health and Human Development.

Public health programs that encourage smoking reduction for pregnant women offer the possibility of effecting a more immediate decline in SIDS. It remains to be determined, however, whether the benefits of public health programs that encourage smoking reduction will include a decline in SIDS occurrence. □

Acknowledgments

An abstract of this research was presented at the April 1991 joint annual meeting of the American Pediatric Society and the Society for Pediatrics Research in New Orleans, La.

We would like to thank Garland Land and his staff at the Missouri Department of Health for making vital statistics data available for this research study. Also, we would like to acknowledge the principal investigators and their staffs, including the scientific advisors and consultants, of the National Institute of Child Health and Human Development SIDS Cooperative Epidemiological Study.

References

1. Damus K, Pakter J, Drongrad E, Standfast SJ, Hoffman HJ. Postnatal medical and epidemiological risk factors for the sudden infant death syndrome. In: Harper RM, Hoffman HJ, eds. *Sudden Infant Death Syndrome: Risk Factors and Basic Mechanisms.* New York, NY: PMA Publishing Corp; 1988:187–201.
2. Naeye RL, Ladis B, Drage JS. Sudden infant death syndrome: a prospective study. *AJDC.* 1976;130:1207–1210.
3. Bergman AB, Wiesner LA. Relationship of passive cigarette smoking to sudden infant death syndrome. *Pediatrics.* 1976;58:665–668.
4. Malloy MH, Kleinman JC, Land GH, Schramm WF. The association of maternal smoking with age and cause of infant death. *Am J Epidemiol.* 1988;128:46–55.
5. Haglund B, Cnattingius S. Cigarette smoking as a risk factor for sudden infant death syndrome: a population-based study. *Am J Public Health.* 1990;80:29–32.
6. Hoffman HJ, Damus K, Hillman L, Krongrad E. Risk factors for SIDS: results of the National Institute of Child Health and Human Development SIDS Cooperative Epidemiological Study. *Ann NY Acad Sci.* 1988;533:13–30.
7. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods.* Belmont, Calif: Lifetime Learning Publications; 1982.
8. SAS Institute Inc. *SAS User's Guide: Statistics, Version 5 Edition.* Cary, NC: SAS Institute Inc; 1985.

Call for Historical Papers on Public Health Nursing

Because in 1993 public health nurses celebrate their centennial, historians of public health nursing are invited to submit articles for review for centennial publications planned by the journal *Public Health Nursing.* Guidelines for authors are printed in the back cover of the journal. Issues are published in March, June, September, and December, and deadlines are 6 months prior to publication. For additional information, please call Karen Buhler-Wilkerson, historical editor, *Public Health Nursing,* at (215) 898-4725.

Stage at Diagnosis in Breast Cancer: Race and Socioeconomic Factors

Barbara L. Wells, PhD, and John W. Horm, MSc

ABSTRACT

Cancer incidence data from three US metropolitan areas were coupled with census tract indicators of education and income. The data suggest that both Black and White cancer patients living in census tracts with lower median education/income values are diagnosed in later disease stages than are patients in tracts with higher median education/income values. Within education and income strata, Black women had a less favorable stage of disease at diagnosis than Whites. The exception was in upper education/income levels, where the disadvantage for Blacks disappeared. These data provide additional evidence that women of low socioeconomic status could benefit from targeted screening. (*Am J Public Health.* 1992;82:1383–1385)

Introduction

Breast cancer has the highest incidence of any cancers among both Black and White women,[1] with about 181 000 new cases expected in 1992.[2] Although the incidence is higher among White women, breast cancer mortality is substantially greater among Black women,[3,4] reflecting their poorer survival experience.[3,5] Survival has differed by income[6] and socioeconomic status level, even when stage adjusted.[7] Stage at diagnosis strongly influences breast cancer survival and, in the few studies investigating socioeconomic status *and* stage at diagnosis, women of lower socioeconomic status with breast cancer were more likely to be diagnosed at a later stage.[8] In comparisons between Black and White women, proportional differences of metastatic breast cancer have been the greatest with the largest differences in income.[9] When Black–White differences in socioeconomic status are small, there are small differences in the distribution of stage at diagnosis.[10] We extend these smaller studies with data from a large population-based set of tumor registry data to investigate the effect of socioeconomic status on Black–White differences in stage of disease at diagnosis and the impact that levels of socioeconomic status have on stage distribution within race.

Materials and Methods

This study used the population-based incidence data, including stage of disease at diagnosis, from the US National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. The SEER Program collects information on all cases of cancer diagnosed among geographically defined subsets of the US population.[1,11] Median values for education and family income at the census tract level were obtained from the 1980 US census and matched with SEER data. Cancers diagnosed from 1978 to 1982 were included to center around the census. The San Francisco–Oakland, Detroit, and Atlanta metropolitan areas were used for two reasons: they are census tracted and therefore could be matched with census tract indicators of socioeconomic status, and they have sufficiently large numbers of Blacks for analysis.

There were 18 113 White and 2861 Black cases of invasive breast cancer reported in the three areas during our study period that could be matched with a census tract (2.1% of the total number of breast cancers could not be matched). Of these, 187 cases were incidental diagnoses at autopsy or the only evidence of their breast cancer was the death certificate with no indication of a prior diagnosis. After excluding these cases, 17 949 invasive breast cancer and 917 in situ cases among White women, and 2838 invasive and 126 in situ cases among Black women, were available for analysis.

The median family income strata were chosen so that each had enough persons of each race for meaningful analysis; $15 000 also was approximately twice the poverty level for a family of four in 1979.[12] Race-specific indicators of median education and family income were used because Whites and Blacks living in the same census tract may have different education and/or income distributions.

The in situ and localized disease stages used here are defined as confined to the site/organ of origin, with the former not being able to metastasize. Cases spread by direct extension to adjacent organs/structures or to regional lymph nodes are considered regional stage. Distant stage includes disease cases involving distant organs or lymph nodes either by direct extension or through discontinuous metastasis.[13] The "other" stage contains those cases that could not be staged from the information contained in medical records or were otherwise unclassified.

Barbara L. Wells is with the Biometry Branch, Division of Cancer Prevention and Control, National Cancer Institute, Bethesda, MD. John W. Horm is with the Division of Health Interview Statistics, National Center for Health Statistics, Hyattsville, Md.

Requests for reprints should be sent to Barbara L. Wells, PhD, Biometry Branch, National Cancer Institute, NIH, EPN-344, 9000 Rockville Pike, Bethesda, MD 20892.

This paper was submitted to the Journal August 26, 1991, and accepted with revisions May 12, 1992.

TABLE 1—Stage Distribution of Breast Cancers among Women, by Race and Census Tract Indicators of Education, Three Areas,[a] 1978 to 1982

	In Situ		Localized		Regional		Distant		Other		Total	
	No.	(%)	No.	(%)	No.	(%)	No.	(%)	No.	(%)	No.	(%)
Under 12 y												
Black	48	(3.9)	466	(37.2)	529	(42.3)	180	(14.4)	29	(2.3)	1252	(100.0)
White	85	(4.3)	902	(45.7)	770	(39.0)	171	(8.7)	47	(2.4)	1890	(100.0)
High school												
Black	64	(4.3)	654	(43.8)	590	(39.5)	142	(9.5)	42	(2.8)	1492	(100.0)
White	455	(4.6)	4583	(46.0)	4007	(40.2)	694	(7.0)	221	(2.2)	9960	(100.0)
13 y and over												
Black	14	(6.4)	110	(50.0)	80	(36.4)	9	(4.1)	7	(3.2)	220	(100.0)
White	377	(5.4)	3429	(49.5)	2623	(37.8)	371	(5.4)	131	(1.9)	6931	(100.0)
Total												
Black	126	(4.2)	1230	(41.5)	1199	(40.4)	331	(11.2)	78	(2.6)	2964	(100.0)
White	917	(4.9)	8914	(47.2)	7400	(39.2)	1236	(6.6)	399	(2.1)	18 866	(100.0)

Note. The results of the Spearman rank-correlation test of association between education and stage of disease were as follows: Black, $t = -5.92$, $P < .001$, and White, $t = -6.32$, $P < .001$.
[a]The three areas are San Francisco–Oakland, Detroit, and Atlanta.

TABLE 2—Stage Distribution of Breast Cancers among Women, by Race and Census Tract Indicators of Income, Three Areas,[a] 1978 to 1982

	In Situ		Localized		Regional		Distant		Other		Total	
	No.	(%)	No.	(%)	No.	(%)	No.	(%)	No.	(%)	No.	(%)
<$15 000												
Black	73	(4.1)	702	(39.8)	720	(40.8)	230	(13.0)	40	(2.3)	1765	(100.0)
White	88	(4.1)	968	(45.4)	850	(39.8)	171	(8.0)	56	(2.6)	2133	(100.0)
$15 000–$24 999												
Black	38	(3.1)	417	(42.9)	399	(41.0)	92	(9.5)	27	(2.8)	973	(100.0)
White	345	(5.7)	3425	(47.0)	2860	(39.2)	528	(7.2)	135	(1.9)	7293	(100.0)
$25 000+												
Black	15	(6.6)	111	(49.1)	80	(35.4)	9	(4.0)	11	(4.9)	226	(100.0)
White	484	(5.1)	4521	(47.9)	3690	(39.1)	537	(5.7)	208	(2.2)	9440	(100.0)
Total												
Black	126	(4.2)	1230	(41.5)	1199	(40.5)	331	(11.2)	78	(2.6)	2964	(100.0)
White	917	(4.9)	8914	(47.2)	7400	(39.2)	1236	(6.6)	399	(2.1)	18 866	(100.0)

Note. The results of the Spearman rank-correlation test of association between education and stage of disease were as follows: Black, $t = -4.02$, $P < .001$, and White, $t = -3.86$, $P < .001$.
[a]The three areas are San Francisco–Oakland, Detroit, and Atlanta.

To test the relationship between stage of disease at diagnosis and income/education, the Spearman rank-correlation coefficient was computed.[14]

Results

Our data indicate that Black women with breast cancer are diagnosed at a later stage of disease than are White women. Only 45.7% of Black women diagnosed in the more curable localized and in situ stages compared with 52.1% of White women (Table 1). Over 11% of the Black women and less than 7% of Whites had breast cancers that had already metastasized beyond localized or regional into distant stage.

Among both Black and White women, there is a tendency toward a less advanced stage with increasing educational level or income ($P < .001$) (Tables 1 and 2). The percentage of cases with distant disease decreases with increasing educational level or income among women of both racial groups.

Black women in both of the two lowest educational strata and income strata had a poorer stage of disease at presentation than did White women. This pattern was not present in either the upper educational strata or the upper income strata. In each of these two strata, Black women had a slightly better stage of disease at diagnosis than did White women.

The incidence of breast cancer is higher among both Black and White women in the upper education and income categories (Tables 3 and 4). This appears largely due to higher rates in the in situ and localized stages. In the distant stage, however, the pattern is the opposite for both races: incidence decreases with increasing education and income. The differences between educational and income strata are more striking among Black women.

Discussion

The association between female breast cancer stage at diagnosis and socioeconomic status gives credence that those of lower income/education are less likely to participate in breast cancer early detection. This lack of participation is further supported by 1987 National Health Interview Survey findings.[4,15] People in lower income/education strata were less likely to have had a screening mammogram within the past year (9.2% with incomes of less than $10 000, 20.9% with incomes of more than $35 000) and were less likely to have heard of the test (24.6% vs 5.9%). Thus, it appears that the most immediate and practical implication of this investigation's results is that cancer screening programs are especially appropriate in populations of lower socioeconomic status, particularly among Blacks.

Offering breast cancer screening programs to those of lower socioeconomic status and Blacks is not without challenges, however. Women of lower socioeconomic status have less access to preventive and medical care, and asymptomatic persons of lower socioeconomic status may not make optimal use of preventive health services.[16–19]. Among Blacks, there may be even less use of the medical system, as compared with Whites.[20] Office visit and screening procedure costs are reported barriers by many Black women,[21] as are physicians' failure to discuss mammography with women and women's perceptions of breast cancer and screening.[22] This complex assortment of challenges requires that early detection barriers of a population be precisely identified to design and maximize effective interventions. □

TABLE 3—Age-Adjusted (1970 US Standard) Incidence Rates of Breast Cancers per 100 000 Women, by Stage, Education, and Race, Three Areas,[a] 1978 to 1982

	In Situ Rate	Localized Rate	Regional Rate	Distant Rate	Other Rate	Total Rate
Under 12 y						
Black	2.5	26.1	29.0	9.8	1.3	68.7
White	4.6	36.8	32.6	7.5	2.1	83.6
High school						
Black	3.3	33.6	30.4	7.4	2.3	76.9
White	4.4	42.0	36.8	6.1	1.8	91.1
13 y and over						
Black	4.9	42.3	28.9	4.1	2.6	82.7
White	6.0	52.5	40.5	5.7	1.9	106.6
Total						
Black	3.2	31.2	30.1	8.2	2.0	74.7
White	4.9	45.4	37.8	6.1	1.8	96.1

[a]The three areas are San Francisco–Oakland, Detroit, and Atlanta.

TABLE 4—Age-Adjusted (1970 US Standard) Incidence Rates of Breast Cancers per 100 000 Women, by Stage, Income, and Race, Three Areas,[a] 1978 to 1982

	In Situ Rate	Localized Rate	Regional Rate	Distant Rate	Other Rate	Total Rate
<$15 000						
Black	3.1	29.6	29.2	8.8	1.6	72.4
White	3.8	40.8	35.7	9.0	1.3	90.6
$15 000–$24 999						
Black	2.6	31.6	31.4	8.1	2.0	75.8
White	4.7	42.8	36.4	6.3	1.5	91.8
$25 000+						
Black	5.6	41.1	26.4	4.6	4.4	82.0
White	5.2	47.7	39.2	5.6	2.1	99.8
Total						
Black	3.2	31.2	30.1	8.2	2.0	74.7
White	4.9	45.4	37.8	6.1	1.8	96.0

[a]The three areas are San Francisco–Oakland, Detroit, and Atlanta.

Acknowledgments

We thank Dr. Charles Brown for his statistical advice and Dr. Susan DeVesa for her review and suggestions.

References

1. Young JL Jr, Percy CL, Asire AJ, eds. Surveillance, epidemiology, end results: incidence and mortality data, 1973–77. *Natl Cancer Inst Monogr.* 1981;57:1–1082.
2. Boring CC, Squires TS, Tong T. Cancer statistics, 1992. *Cancer.* 1992;42:19–38.
3. Baquet CR, Ringen K, Pollack ES, et al. *Cancer among Blacks and Other Minorities: Statistical Profiles.* Bethesda, Md: National Institutes of Health; 1986. DHEW publication (NIH) 86-2785.
4. *Cancer Statistics Review 1973–1986 Including a Report on the Status of Cancer Control.* Bethesda, Md: National Cancer Institute; 1989. NIH publication 89-2789.
5. Young JL Jr, Ries LG, Pollack ES. Cancer patient survival among ethnic groups in the United States. *JNCI.* 1984;73:341–352.
6. Lipworth L, Abelin T, Connelly RR. Socioeconomic factors in the prognosis of cancer patients. *J Chronic Dis.* 1970;23: 105–115.
7. Dayal HH, Power RN, Chiu C. Race and socio-economic status in survival from breast cancer. *J Chronic Dis.* 1982;35:675–683.
8. Farley TA, Flannery JT. Late-stage diagnosis of breast cancer in women of lower socioeconomic status: public health implications. *Am J Public Health.* 1989;79: 1508–1512.
9. Polednak AP. Breast cancer in Black and White women in New York State: case distribution and incidence rates by clinical stage at diagnosis. *Cancer.* 1986;58:807–815.
10. Polednak AP. A comparison of survival of Black and White female breast cancer cases in upstate New York. *Cancer Detect Prev.* 1988;11:245–249.
11. Horm JW, Asire AJ, Young JL Jr, Pollack ES, eds. *SEER Program: Cancer Incidence and Mortality in the United States 1973–81.* Bethesda, Md: National Institutes of Health; 1984. NIH publication 85-1837.
12. Bureau of the Census. *1980 Census of the Population: Detailed Population Characteristics, Part 1 United States Summary.* Washington, DC: US Department of Commerce; 1984. PC80-1-D1-B.
13. *1987 Annual Cancer Statistics Review Including Cancer Trends: 1950–1985.* Bethesda, Md: National Cancer Institute; 1988. NIH publication 88-2789.
14. Siegel S. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill; 1956.
15. *Vital and Health Statistics. Breast Cancer Risk Factors and Screening: United States, 1987.* Hyattsville, Md: National Center for Health Statistics; 1990. DHHS publication 90-1500.
16. Maclean U, Sinfield D, Klein S, Harnden B. Women who decline breast screening. *J Epidemiol Community Health.* 1984;38: 278–283.
17. French K, Porter AMD, Robinson SE, McCallum FM, Howie JG, Roberts MM. Attendance at a breast screening clinic: a problem of administration or attitudes. *Br Med J.* 1982;285:617–620.
18. Vernon SW, Laville EA, Jackson GL. Participation in breast screening programs: a review. *Soc Sci Med.* 1990;30:1107–1118.
19. Roberts MM, Alexander FE, Elton RA, Rodgers A. Breast cancer stage, social class and the impact of screening. *Eur J Surg Oncol.* 1990;16:18–21.
20. Blendon RJ, Aiken LH, Freeman HE, Corey CR. Access to medical care for Black and White Americans. *JAMA.* 1989;261: 278–281.
21. Burack RC, Liang J. The acceptance and completion of mammography by older Black women. *Am J Public Health.* 1989; 79:721–726.
22. Fox SA, Stein JA. The effect of physician-patient communication on mammography utilization by different ethnic groups. *Med Care.* 1991;29:1065–1082.

Measuring the Use of Mammography: Two Methods Compared

Darrah Degnan, BSc, Russell Harris, MD, MPH, Jane Ranney, PhD, Dana Quade, PhD, Jo Anne Earp, ScD, and Jorge Gonzalez, MD

ABSTRACT

Population studies often estimate mammography use using women's self-reports. In one North Carolina county, we compared self-report surveys with a second method—counting mammograms per population—for 1987 and 1989. Estimates from self-reports (35% in 1987, 55% in 1989) were considerably higher than those from mammogram counts (20% in 1987, 36% in 1989). We then confirmed 66% of self-reports in the past year. Self-reported use is more accurate regarding whether a woman has had a mammogram than when she had it, but self-reports accurately measure change over time. (*Am J Public Health.* 1992;82:1386–1388)

Introduction

A number of studies have examined the percentage of women in the population receiving mammograms.[1–4] Most of these studies rely on self-reported behavior, subject to the errors of this method.[5–7]

We collected self-report data on mammography use by a random-digit dialing telephone survey in a single county of North Carolina.[2] Using institutional records, we also documented mammograms performed on county residents for 1987 and 1989. This paper compares these two methods of measuring mammography use.

Methods

This study was approved by the Committee on the Protection of Human Subjects of the University of North Carolina School of Medicine.

Site and Population Characteristics

The total population of New Hanover County, NC, is 120 000, with about 14 000 women aged 50 to 74 years.[8] Because New Hanover contains a regional medical center and is geographically isolated, nearly all county residents use local sources of medical care.

Data Sources

Telephone survey self-reports. In January 1988 and 1990, we conducted 15-minute random-digit dialing telephone interviews with county women aged 50 to 74 years. Women were asked, "Have you ever had a mammogram?" and "How long ago did you have your last mammogram?" The telephone response rates were 73% (n = 487) in 1988 and 74% (n = 486) in 1990. Respondents were invited to give their names for follow-up; 85% did so.

Mammogram counts from institutional records. From all radiology centers, we obtained logbook or computer report listings of all county women who had had mammograms in 1987 and 1989, and we collected a copy of the mammogram reports. Data (name, age, social security number, zip code, date of mammogram) from all centers were combined into a single database. Women not aged 50 to 74 years and out-of-county residents were excluded. We then divided the number of county women who had had a mammogram by the number of all county women aged 50 to 74, using population estimates from the North Carolina Office of State Management and Budget.

Validating January 1990 Self-Reports against Institutional Records

We supplemented the 1987/1989 mammogram listings with listings from 1988 and January 1990. We matched "self-report" women by date of birth and name with their most recent mammogram in the expanded database. We confirmed matches by comparing the self-reported telephone numbers with those on the mammogram reports. To verify the identity of women not matched, we used telephone numbers from public directories. We also searched institutional databases for women not matched, both those reporting and those denying a past-year mammogram. We did not attempt to verify reports of women who claimed mammography outside the county (n = 19).

Analysis

We calculated the proportion of women who reported having had a mammogram for 1987 and 1989. We used linear regression to estimate the relationship between reported months and actual months

Darrah Degnan, Russell Harris, and Jo Anne Earp are with the Lineberger Comprehensive Cancer Center at the University of North Carolina (UNC) at Chapel Hill, NC; Dr. Harris is also with the Center for Health Promotion and Disease Prevention at UNC. Dr. Harris and Jorge Gonzalez are with the UNC School of Medicine; Dr. Earp, Dr. Harris, and Dana Quade are with the UNC School of Public Health; and Dr. Gonzalez and Jane Ranney are with the Wilmington, NC, Area Health Education Center.

Requests for reprints should be sent to Darrah Degnan, Department of Epidemiology, CB 7400, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7400.

This paper was submitted to the Journal September 24, 1991, and accepted with revisions March 24, 1992.

TABLE 1—Verification of Mammogram Self-Reports of Women Aged 50 to 74 Years in New Hanover County, North Carolina, 1989

Self-Reported Date of Most Recent Mammogram	Actual Date of Mammogram Confirmed by Mammogram Report: Within Past Year	More Than 1 Year Ago; Not More Than 2 Years Ago	More Than 2 Years Ago	No Report Found	Total
Within past year	164	58	8	17[a]	247
More than 1 year ago but not more than 2 years ago	1	22	11	5	39
More than 2 years ago	0	0	5	12	17
None	1	0	0	152	153
Total	166	80	24	186	456

[a]Includes two who had chest X-rays and eight whose identity could not be independently verified.

since the last mammogram. We used the χ^2 statistic to relate the accuracy of the self-report to age, education, employment, and income groups.

Results

In January 1988, 35% of interviewed women reported having had a mammogram in the past year; this increased to 55% in 1990.[9] The mammogram counts from institutional records were 2650 (20% of the population of 13 501) for 1987 and 4994 (36% of population of 14 025) for 1989.[9] After excluding 19 women who reported a past-year mammogram but outside the county, we matched mammogram records in 1989 for 164 of the remaining 247 women (66%) interviewed in January 1990 who reported a past-year mammogram in the county (Table 1). The 164 confirmed self-reports totaled 34% of the initial survey sample (n = 486), compared with 36% of the same population found by counting mammograms (Figure 1).

Incorrect recollection of the date—actually longer than past-year—was more common (66/247 or 27%) than unconfirmed reports (17/247 or 7%). The regression line equation for plotting reported time since last mammogram against actual time was actual time = 2.54 + (1.02 × reported time), indicating that actual time was about 3 months longer on average than reported time. Among those denying a mammogram, one report was found. There was no statistically significant association between accuracy of self-report and age, education, employment, or income.

Discussion

Women aged 50 to 74 in one county reported higher mammography usage within the previous year than that calculated from institutional records (35% vs 20% in 1987; 55% vs 36% in 1989). Change over time was similar by both methods (20% vs 16%). Of those interviewed in 1990 who reported having had a mammogram within the county in the past year, 66% were confirmed by mammogram records. We found the overreporting by self-report was due primarily to underestimating the time since the last mammogram ("telescoping"), not to unconfirmed usage.

Although telephone interviews do not gather information about women without telephones (12% of all households) or about nonrespondents, the response rate of our survey was acceptable for a population survey,[5] and the percentage of women with validated self-reports (34%) agreed closely with usage calculated from institutional records (36%).

To our knowledge, no previous study has examined mammogram counts per population as a method of measuring mammography use. Although the mammogram count method avoids the overreporting found in the self-report method, it introduces a potential undercount if all mammograms are not located. We are confident that we located nearly all mammograms because (1) the setting made it likely that nearly all county residents would have had mammography within the county, and (2) we were able to document a mammogram of some date on all but 7% of the respondents.

Our results agree with those of Gordon et al. (personal communication, October 23, 1991), who confirmed 67% of

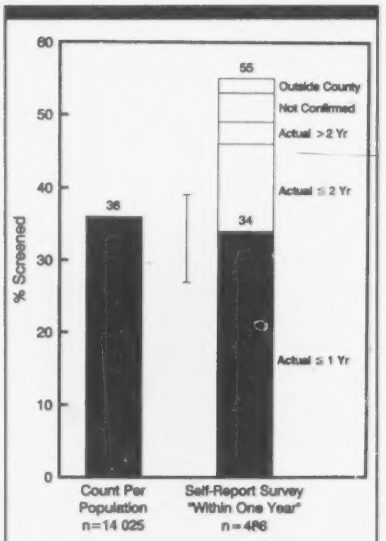


FIGURE 1—Percentage of women aged 50 to 74 screened in past year, New Hanover County, North Carolina, 1989.

reported mammograms within the previous year for women aged 50 to 74 years in a large health maintenance organization. They differ, however, from those of King et al.,[10] who verified 94% in another health maintenance organization where respondents were concurrently cued to get mammography. We agree with others[7,11,12] that the telescoping phenomenon should be taken into account in survey research.

Measuring mammography rates in a general population by self-report is accurate for determining whether a woman has had a mammogram but is less accurate for determining when. Surveys that ask "when was your last mammogram" will overestimate usage. Such surveys can, however, accurately estimate change in usage. In suitable areas, researchers can estimate usage directly by counting institutional records. □

Acknowledgments

This research was supported in part by National Cancer Institute Grant R01CA45022.

We wish to thank the staff of Delaney Radiology, New Hanover Regional Medical Center, Cape Fear Memorial Hospital, and Wilmington Imaging for their generous assistance and support.

References

1. Anda RF, Sienko DG, Remington PL, Gentry EM, Marks JS. Screening mammography for women 50 years and older: current practices and trends, 1987. *Am J Prev Med.* 1990;6:123–129.

2. NCI Breast Cancer Screening Consortium. Screening mammography: a missed clinical opportunity? results of the NCI Breast Cancer Screening Consortium and National Health Interview Survey Studies. *JAMA.* 1990;264:54–58.
3. National Cancer Institute: Use of mammography—United States 1990. *MMWR.* 1990;39:621, 627–630.
4. Use of mammography for breast cancer screening—Rhode Island, 1987. *MMWR.* 1987;37:357–360.
5. Mathiowetz NA. Validity of reporting in surveys. In: Fowler FJ, ed. *Proceedings of the 5th Conference on Health Survey Research Methods.* Rockville, Md: National Center for Health Research; 1989:91–95. DHHS publication PHS 89–3447.
6. Cannell CF, Marquis KH, Laurent A. A summary of studies of interviewing methodology. *Vital Health Stat* [2]. 1977; no. 69. DHEW publication HRA 77–1343.
7. Sudman S, Bradburn NM. Effects of time and memory factors on response in surveys. *J Am Stat Assoc.* 1973;68:805–815.
8. 1989 population data (extrapolated annually from 1980 census data). Raleigh, NC: North Carolina Office of State Management and Budget.
9. Harris R, Fletcher S, Gonzalez J, et al. The growing racial gap in mammography. Presented at the 14th Annual Meeting of the Society of General Internal Medicine; May 1991. Abstract.
10. King ES, Rimer BK, Trock B, Balshem A, Engstrom P. How valid are mammography self-reports? *Am J Public Health.* 1990;80:1386–1388.
11. Sawyer JA, Earp JA, Fletcher RH, Daye FF, Wynn TM. Accuracy of women's self-report of their last Pap smear. *Am J Public Health.* 1989;79:1036–1037.
12. Gordon NP, Hiatt RA, Lampert DI. Obtaining information about cancer screening histories: a comparison of questionnaire data with medical record documentation. Presented at the 118th Annual Meeting of the American Public Health Association; September 30–October 4, 1990; New York, NY. Abstract.

Women with Multiple Sexual Partners: United States, 1988

Stuart N. Seidman, MD, William D. Mosher, PhD, and Sevgi O. Aral, PhD

ABSTRACT

Women who have multiple sexual partners in a short time period are appropriate targets for sexually transmitted disease (STD) prevention. We analyzed survey data collected in 1988 from a nationally representative sample of 8450 American women aged 15 to 44 to identify markers of such behavior. Among sexually active persons, 0.4% of married women and 8.4% of unmarried women had two or more sexual partners in the 3 months preceding the interview; unmarried marital status, early age at first sexual intercourse, lack of religious affiliation, and young age were associated with this behavior. All except young age were predictive after multivariate analysis. Such factors may help define women at elevated STD risk and allow better targeting of STD prevention. (*Am J Public Health.* 1992;82:1388–1394)

Introduction

Each additional sexual partner puts an individual at some incremental risk of exposure to a sexually transmitted pathogen and increases the dissemination of such pathogens. The number of lifetime sex partners quantifies an individual's cumulative risk for persistent viral sexually transmitted diseases (STDs) and cervical cancer[1–5] and is associated with bacterial STDs, including pelvic inflammatory disease.[6–8] The number of sex partners in a recent time interval, such as the previous several months, is a more precise measure for assessing the risk for incident infection: among STD clinic patients, having multiple partners in a specific recent time period (less than 6 months) is a risk factor for the acquisition of bacterial STDs, such as gonorrhea,[9,10] chlamydia,[11–13] syphilis,[14] and chancroid,[15] as well as for the presence of viral STDs.[16–18] Furthermore, for STDs with a short period of infectiousness, such as gonorrhea, having multiple partners in a short time span increases the rate of dissemination in the population.

One way to identify individuals at greatest risk is to focus on those who report having two or more partners in a specific time interval. Having multiple recent partners is associated with disease risk for at least two reasons: first, it reflects the increased likelihood of encountering a sexually transmitted pathogen through having multiple potential exposures, and second, it may reflect an increased probability of choosing a partner with an infection through a riskier pattern of partner recruitment.[19,20] Primary prevention strategies to reduce the number of sexually transmitted infections can use the characteristics of these women to target education and counseling that encourage safe sexual practices. Secondary prevention strategies, aimed at early diagnosis and treatment of asymptomatic infection, can target these high-risk women for selective screening (e.g., for human immunodeficiency virus, syphilis, and/or pelvic inflammatory disease).

Over this century, Americans have increasingly had greater numbers of sexual relationships, in part because of postponed marriage, increases in separation and divorce, and more reliable contraception.[21,22] A growing majority of unmarried

Stuart N. Seidman and Sevgi O. Aral are with the Division of STD/HIV Prevention, National Center for Prevention Services, and William D. Mosher is with the National Survey of Family Growth, National Center for Health Statistics, Centers for Disease Control, Atlanta, Ga.

Requests for reprints should be sent to Sevgi Aral, PhD, Behavioral Epidemiologist, Division of STD/HIV Prevention, Centers for Disease Control MS-E02, 1600 Clifton Rd, Atlanta, GA 30333.

This paper was submitted to the Journal September 4, 1991, and accepted with revisions April 9, 1992.

women are sexually active,[23,24] and many young women who reduce the number of sexual partners they have because of fear of acquired immunodeficiency syndrome (AIDS) continue to have more than one partner.[25] Overall, almost two thirds of sexually experienced women report having had two or more sexual partners during their lifetime,[23,24] and this proportion is increasing. These developments are likely to portend even greater proportions of sexually active women, especially young women, having multiple sexual partners.

In the 1988 through 1990 General Social Survey, a nationally representative survey that included 1123 women 18 to 49 years of age, 13% of these women reported having two or more sexual partners in the preceding year (General Social Survey, unpublished data). A combined analysis of this survey and one conducted by the *Los Angeles Times* suggests that about one fourth of unmarried women of reproductive age had two or more sexual partners in the preceding year.[21] From the General Social Survey, factors associated with a greater mean number of partners in the preceding year were young age, single marital status, and central city residence.[26] Similarly, in a self-administered questionnaire from clients of a family planning clinic in Pennsylvania, factors that were associated with women reporting five or more sexual partners in the previous year included young age, unmarried marital status, and lack of any college education.[19]

Each of the previous studies of women with multiple partners has added significantly to the demographic description of these women. However, interpretation of data from these studies is somewhat limited by methodological concerns. For example, the *Los Angeles Times* survey had a response rate of only 33%.[21] Most significantly, none addressed the prevalence of having multiple partners in a recent short time period.

The present study uses face-to-face interviews of a representative sample of American women to focus on factors associated with having multiple partners in the 3 months preceding the interview.

Methods

We obtained data for this analysis from the National Survey of Family Growth Cycle IV conducted by the National Center for Health Statistics. The National Survey of Family Growth is based on personal interviews with women of reproductive age regarding fertility, family planning, and maternal and child health. The Cycle IV survey was conducted between January and August 1988 with a multistage probability sample of 8450 women 15 to 44 years of age in the civilian noninstitutionalized population of the United States. The sample included 5354 White women, 2771 Black women, and 325 women of other races. In personal interviews lasting just over an hour, respondents provided detailed information on their reproductive health, including age at first intercourse, and number of recent sexual partners. Response rates were high overall (82.5%), and similar for Black (82.2%) and White (82.6%) women. For most of this analysis, only women who reported having had sexual intercourse in the preceding year were included. This reduced the sample size to 7011 women.

Reported percentages are weighted national estimates. The weight for each respondent represents her probability of selection for the survey, adjusted for nonresponse and adjusted again to agree with US Bureau of the Census estimates by age, race, marital status, and parity.[27] The *z* tests used in Tables 1 through 3 applied the formulas and procedures recommended for this survey.[27] These formulas are based on the balanced half-sample replication technique, an accurate method for estimating sampling errors from complex samples such as this one. These estimates take into account both the sample size and sample design on which our weighted percentages are based.

Near the end of the interview, after questions about pregnancy, contraception, and marriage, all women who had ever had intercourse were asked: "In the last 3 months in which you were having intercourse, with how many men did you have intercourse?" The variables and categories shown here were arrived at after preliminary analysis, using both cross tabulation and logistic regression. Our tables show women who report two or more partners in the preceding 3 months as a percentage of all women in a particular group who had intercourse in the previous year. Relative risks and confidence intervals were determined from these proportions for each group. To determine which variables were independently associated with having multiple recent partners when the effects of other variables were controlled, we used multiple logistic regression.

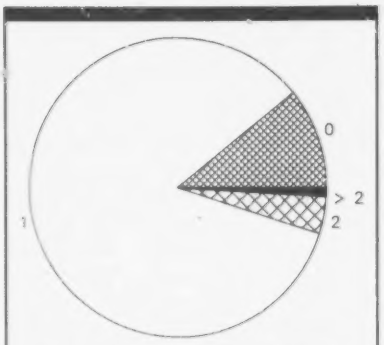


FIGURE 1—Number of recent (previous 3 months) sex partners among women 15 to 44 years of age.

Results

Overall, in 1988, 11% of all women of reproductive age did not have intercourse at all in the preceding 3 months, 85% had intercourse with one man during that period, 4% had intercourse with two men during that period, and less than 1% report having intercourse with three or more men in the preceding 3 months (Figure 1).

Of sexually active women (i.e., those who have had sexual intercourse in the preceding year), 3.4% report having intercourse with two or more men in the previous 3 months (Table 1). Marital status is the most important predictor of having two or more partners: 0.4% of married women report multiple partners in a 3-month period compared with 7.9% of never-married women and 9.5% of divorced/separated women (Table 1). Since the proportion having multiple partners is small among married women and substantially higher among never-married and divorced/separated women, and since STD incidence among Blacks and Whites differs dramatically, the results are stratified by marital status and race.

Among all sexually active women, increasing age is associated with lower proportions reporting multiple recent partners (Table 1). This is related to the greater proportion of married women in the older age groups. Among those who are not married, however, the pattern among Blacks reveals a later peak in this behavior among women in their thirties (Figure 2). Among unmarried, sexually active White women, most of those who report having two or more recent partners are under 30 years of age; the unmarried, sexually active Black women who report this most are those 30 or older. In fact, further stratification reveals that 35- to 39-year-old

TABLE 1—Sexually Active[a] Women 15 to 44 Years of Age Reporting Two or More Partners in the Previous 3 Months of Sexual Activity: United States, 1988

	n	Women Reporting ≥ 2 Partners, %
Marital status[b]		
Married	3971	0.4**
Unmarried	2992	8.4
Never married	2037	7.9
Divorced/separated	955	9.5
Race		
Black	2341	5.2*
White	4428	3.1
Other	242	3.6
Age, y		
15–19	638	8.2
20–24	1083	5.4
25–29	1433	3.4
30–34	1568	2.0
35–39	1270	2.1
40–44	1019	2.0
Family income/poverty		
≤1.5 × poverty line	1954	6.5**
>1.5 × poverty line	5057	2.6
Age at first sexual intercourse[c]		
Under 15	735	9.3*
15–17	3019	4.3*
18+	3252	1.8*
Religious affiliation		
Has an affiliation	6581	3.1**
Has no affiliation	430	8.7
Total	7011	3.4

[a]In this and subsequent tables, women who had intercourse in the 12 months before they were interviewed.
[b]Total includes 48 widowed women who are not shown separately.
[c]Women 20–44 years of age.
*Difference between this category and the category below it is significant at the .05 level using a two-tailed *t* test.
**Difference between this category and the category below it is significant at the .01 level using a two-tailed *t* test.

divorced/separated Black women, in particular, report two or more recent partners more often than any other race–marital status group (16.4%).

Poverty status is associated with distinct patterns of this behavior among women of reproductive age overall (Table 1), but not among unmarried women in particular. In general, less education is associated with multiple recent partners among all unmarried women except divorced/separated White women. Never-married White women with a high school diploma only are significantly more likely to report multiple recent partners than those with some college education or no high school diploma (relative risk = 2.4, $P < .05$) (Table 2). Among never-married Black women, formal education is inversely associated with multiple recent partners: 9.2% of those in the lowest education group report having two or more recent partners compared with 2.7% among those with some college ($P < .05$). In contrast, among divorced/separated White women, educational level is positively associated with having multiple partners: 5.5% of those without a high school diploma, 7.3% of those with only a high school diploma, and 13.7% of those with some college had two or more partners in the preceding 3 months ($P < .05$) (Table 3).

Earlier age at first sexual intercourse is strongly associated with a higher rate of having multiple recent partners for all never-married women except divorced/separated Whites. Among never-married White women, the percentage reporting multiple partners increases almost fivefold for those whose first intercourse was before age 15 (17.8%) compared with those whose first intercourse was after age 18 (3.8%). This association is also large among never-married Black women, substantially smaller among divorced/separated Black women, and nonexistent among divorced/separated White women.

Among never-married women, having two or more recent sex partners is reported most by those with no religious affiliation and least by those who attend religious services once a month or more; there is no such significant association among divorced/separated women. Among never-married Whites, those with no religious affiliation are more likely to have multiple partners than those with some affiliation (16.7% vs 7.4%; $P < .05$). Among never-married Blacks, lack of at least monthly attendance at religious services (compared with monthly attendance) is associated with having multiple recent partners (8.8% vs 4.3%; $P < .05$).

Among never-married White women, residence in the central city of a metropolitan area is associated with a significantly higher percentage reporting multiple recent sex partners (11.3%) than is suburban or nonmetropolitan residence (7.4%, $P < .05$). This is not true among Black women or divorced/separated women.

Multivariate Analysis

To control for the simultaneous effects of age, education, residence in a metropolitan area, age at first sexual intercourse, religious affiliation, and religiosity on the number of recent partners, we conducted multiple logistic regression analyses. Because of interactions among race, marital status, and the other variables, we conducted separate analyses for the four race–marital status groups.

Two key findings emerge from this analysis (Table 4). First, among sexually active women overall, early age at first sexual intercourse predicts a higher likelihood of having two or more sexual partners in the previous 3 months of sexual activity. After stratification by race and marital status, this finding remains significant for never-married women but does not hold for divorced/separated women. Second, religion is associated with having multiple recent sex partners both for women who have never married and for those who are divorced or separated. However, different measures of religion's effect are significant for Whites compared with Blacks: among Whites, religious affiliation independently predicts this behavior, while among Blacks regular attendance at religious services is predictive.

After controlling for the effects of other variables, the significant predictors

of having multiple recent sex partners among White women who have never married are early first sexual intercourse (younger than 18 years compared with 18 or older), central city residence, and having no religious affiliation. Among Black women who have never married, first sexual intercourse before 15 years of age and less than monthly attendance at religious services are predictive.

Independent predictors of having multiple recent partners among divorced/separated White women include any college education and having no religious affiliation. Among divorced/separated Black women, only less than monthly attendance at religious services predicts this behavior.

Discussion

We found that among sexually active women, less than 1 in 200 of those who are currently married and less than 1 in 10 of those who are currently unmarried report having multiple recent sexual partners. Marital status, socioeconomic status, central city residence, age at first sexual intercourse, religious affiliation, and religious attendance are associated with the expression of this sexual behavior. Furthermore, the effect of each of these factors is different among Black compared with White sexually active, unmarried women.

We studied four groups in detail in this analysis: never-married Whites, never-married Blacks, divorced/separated Whites, and divorced/separated Blacks. Race and marital status were singled out for stratification because these are generally accepted as markers of distinct patterns of sexual behavior[28] and because they provide discrete population subgroups for the targeting of STD interventions. Two variables stand out because their effects are strong and because their effects were found in multiple race–marital status groups: age at first sexual intercourse and the religious factors.

We found that earlier first intercourse predicts having multiple recent partners for never-married Whites and Blacks and for divorced/separated Blacks, even when the effects of other variables are controlled. Early first sexual intercourse has been associated with "risky" behaviors such as using drugs,[6,29] not using contraception at first intercourse,[30] having more sexual partners,[31,32] and having more frequent intercourse.[33,34] It may also be a marker for other sexual behaviors that place an individual at increased risk for STDs, such as lack of condom use, less discriminating sex partner recruitment, and having multiple sex partners in a short time period.[35] Finally, early age at first intercourse is directly linked to bacterial and viral STDs[1,2,36,37] and to cervical cancer and dysplasia.[38–41]

The association we have demonstrated between low religiosity and having multiple recent partners among unmarried Blacks, and between lack of religious af-

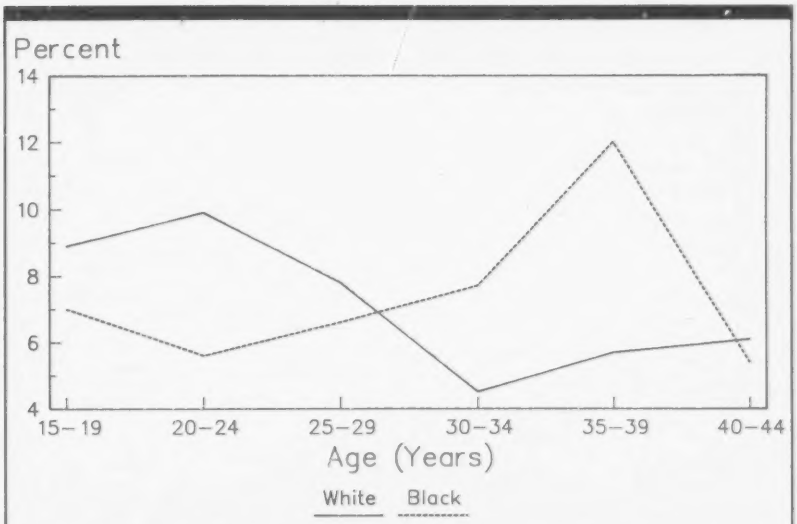


FIGURE 2—Sexually active, unmarried women reporting two or more partners in previous 3 months by age and race.

TABLE 2—Percentage of Never Married Sexually Active White and Black Women Reporting Two or More Sexual Partners in the Previous 3 Months of Sexual Activity: United States, 1988

	White, Relative Risk (%)	Black, Relative Risk (%)
Education, y[a]		
0–11	1.0 (5.3)	3.4* (9.2)
12	2.4* (13.1)	2.3 (6.2)
13+	1.0 (5.5)	1.0 (2.7)
Age at first sexual intercourse		
Under 15	4.7** (17.8)	5.8** (13.9)
15–17	2.5*** (9.4)	2.3* (5.5)
18+	1.0 (3.8)	1.0 (2.4)
Religiosity		
No religious affiliation	3.0** (16.7)	2.3* (10.1)
Some affiliation		
Attends services less than one time per month	1.5 (8.6)	2.0* (8.8)
Attends services more than one time per month	1.0 (5.6)	1.0 (4.3)
Residence		
Central city metro	1.5* (11.3)	1.0 (6.2)
Suburban/nonmetro	1.0 (7.4)	1.0 (6.2)
Total	... (8.3)	... (6.2)

[a]Women 20–44 years of age.
*Significantly different from the reference category (relative risk = 1.0) at the .05 level.
**Significantly different from the reference category (relative risk = 1.0) at the .01 level.
***Significantly different from the reference category (relative risk = 1.0) at the .001 level.

TABLE 3—Percentage of Divorced/Separated Sexually Active White and Black Women Reporting Two or More Sexual Partners in the Previous 3 Months of Sexual Activity: United States, 1988

	White, Relative Risk (%)	Black, Relative Risk (%)
Education, y[a]		
0–11	1.0 (5.5)	1.0 (16.2)
12	1.3 (7.3)	0.5 (8.4)
13+	2.5* (13.7)	0.7 (11.3)
Age at first sexual intercourse		
Under 15	1.0 (8.7)	2.5* (16.8)
15–17	1.1 (9.5)	1.8 (11.8)
18+	1.0 (8.9)	1.0 (6.6)
Religiosity		
No religious affiliation[b]	.....	.....
Some affiliation		
Attends services less than one time per month	1.5 (8.9)	1.0 (11.4)
Attends services more than one time per month	1.0 (6.1)	1.0 (11.3)
Residence		
Central city metro	1.0 (7.2)	1.0 (10.8)
Suburban/nonmetro	1.4 (9.9)	1.1 (11.6)
Total	... (9.2)	... (11.2)

[a]Women 20–44 years of age.
[b]n < 60.
*Significantly different from the reference category (relative risk = 1.0) at the .05 level.

TABLE 4—Factors Associated with Having Two or More Recent Sexual Partners by Marital Status and Race: Results of Logistic Regression

	Odds Ratio	95% CI
Never-married Whites (n = 884)		
First sexual intercourse < 15 years of age (vs after 18 years)	6.6	2.4, 18.8
First sexual intercourse 15–17 years of age (vs after 18 years)	3.6	1.5, 8.8
Central city residence	2.3	1.3, 4.0
Lack of any religious affiliation	2.1	1.1, 4.3
Never-married Blacks (n = 1086)		
First sexual intercourse < 15 years of age (vs after 18 years)	4.4	2.0, 10.1
Less than monthly attendance at religious services	2.1	1.3, 3.4
Divorced/separated Whites (n = 499)		
Any college education	2.1	1.1, 4.0
Lack of any religious affiliation	3.4	1.3, 8.9
Divorced/separated Blacks (n = 441)		
Less than monthly attendance at religious services	2.0	1.0, 3.8

filiation and having multiple recent partners among unmarried Whites, suggests that the religious variables may be markers of women who are more apt to disregard behavioral norms with respect to sexual behavior. Having a religious affiliation and attending religious services frequently have been inversely associated with early first sexual intercourse,[29,42] number of partners in the preceding year (General Social Survey, 1988–1990, unpublished data), and coital frequency.[34] Some theorists view one's connection to religion as an indication of proneness to engage in socially acceptable behavior.[29,34] The relationship between such markers and risky behavior and their potential use in targeting interventions require further research.

In our data, some factors have particularly strong effects in some race–marital status groups, but not in others. For example, early first sexual intercourse and low religiosity, which predict having multiple recent partners among never-married women, are much weaker or unimportant as predictors among divorced/separated women. Educational level is positively associated with having multiple partners among divorced/separated White women, and negatively among never-married women. In addition, divorced/separated Black women are almost twice as likely to have multiple recent partners as are never-married Black women (11% vs 6%; $P < .05$). Thus, being formerly married (compared with never married) is associated with different predictors of having multiple recent partners, and also distinguishes some groups (e.g., Black women) at higher "risk" for having multiple partners.

Similarly, race is associated with different patterns of this behavior among unmarried, sexually active women. The data suggest different, though not necessarily riskier, practices among Black women compared with White women. The finding that having multiple recent sexual partners is most common among divorced/separated Black women (11.2%), least common among never-married Black women (6.2%), and intermediate among unmarried White women (9.0%) points to the oversimplification introduced by using race alone as a predictor of sexual behavior; stratification by marital status adds a new dimension to the discussion of such racial differences. An accurate appraisal of these differences is especially important for STD prevention activities because of the wide disparity in STD incidence between Blacks and Whites.

These self-reported data may contain bias. Intentional nonreporting, incomplete recall, and misunderstanding of the survey questions could lead to inaccuracies in self-reported number of sexual partners and result in underreporting of multiple partners. Thus, these findings probably represent minimum estimates of the number of recent sex partners. However, no representative database exists for comparison.

This analysis focuses on one high-risk behavior for STD acquisition (i.e., multiple current partners) and describes profiles of women who report this behavior and who should be targets for STD interventions. However, the contribution of behavioral factors to a woman's risk for STD acquisition must be seen in the context of the specific pathogen and the prevalence of that infection in her "partner pool." For most groups of women, in fact, STD rates do not parallel the prevalence of having multiple partners. Thus, although having two or more recent sex partners is clearly associated with STDs, factors such as condom use and the prevalence of disease in the partner pool also have important effects on STD acquisition.[43–45] Our data focus attention on the women who are at some elevated risk for STDs as a result of this specific behavior or other risky behaviors (e.g., indiscriminate partner recruitment)[12,20] for which having multiple recent partners may be a marker.

The charge of public health practitioners is to define high-risk behavior, describe high-risk populations, and design and implement appropriate interventions. This description of some characteristics of women who report having multiple recent partners allows practitioners and educators to promote safe sexual practices among certain target groups—such as younger women, unmarried women whose first sexual intercourse was before age 15, never-married White women who live in a city and/or have no religious affiliation, divorced/separated White women who are college educated, and unmarried Black women who did not finish high school. Early intervention focused on those most likely to exhibit such risky behavior can lead to a reduction in lifetime STD acquisition and transmission. Finally, clinical intervention (e.g., screening for asymptomatic STDs, high clinical index of suspicion for STDs) also can be appropriately targeted to reduce STD morbidity. □

Acknowledgments

The National Survey of Family Growth was supported by the National Center for Health Statistics, the Office of Population Affairs, and the National Institute for Child Health and Human Development, all of the US Department of Health and Human Services. These and other agencies also contributed to the design of the survey.

A portion of this research was presented at the annual meeting of the Society of Behavioral Medicine held in Washington, DC, March 1991.

We thank Willard Cates Jr, MD, MPH, Tom Peterman, MD, and William Darrow, PhD, for their critical reviews and insightful comments.

References

1. Kjaer SK, Engholm G, Teisen C, et al. Risk factors for cervical human papillomavirus and herpes simplex virus infections in Greenland and Denmark: a population-based study. *Am J Epidemiol.* 1990;131: 669–682.
2. Stavraky KM, Rawls WE, Chiavetta J, Donner AP, Wanklin JM. Sexual and socioeconomic factors affecting the risk of past infections with herpes simplex virus type 2. *Am J Epidemiol.* 1983;118:109–121.
3. Alter MJ, Ahtone J, Weisfuse I, Starko K, Vacalis TD, Maynard JE. Hepatitis B virus transmission between heterosexuals. *JAMA.* 1986;256:1307–1310.
4. Breinig MK, Kingsley LA, Armstrong JA, et al. Epidemiology of genital herpes in Pittsburgh: serologic, sexual, and racial correlates of apparent and inapparent herpes simplex infections. *J Infect Dis.* 1990;162:299–305.
5. Darrow WW, Echenberg DF, Jaffe HW, et al. Risk factors for human immunodeficiency virus (HIV) infections in homosexual men. *Am J Public Health.* 1987;77:479–483.
6. Li DK, Daling JL, Wang SP, Grayston JT. Evidence that *Chlamydia pneumoniae*, strain TWAR, is not sexually transmitted. *J Infect Dis.* 1989;160:328–331.
7. Washington AE, Aral SO, Wolner-Hanssen P, Grimes DA, Holmes KK. Assessing risk for pelvic inflammatory disease and its sequelae. *JAMA.* 1991;266:2581–2586.
8. Aral SO, Mosher WD, Cates W Jr. Self reported pelvic inflammatory disease in the U.S.: 1988. *JAMA.* 1991;266:2570–2573.
9. Cates W Jr, Wasserheit JN. Gonorrhea, chlamydia, and pelvic inflammatory disease: a review of the recent literature. *Curr Opinions Infect Dis.* 1990;3:10–15.
10. Pederson AHB, Bonin P. Screening females for asymptomatic gonorrhea infection. *Northwest Med.* 1971;70:255–261.
11. Schachter J, Stoner E, Moncada J. Screening for chlamydia infections in women attending family planning clinics. *West J Med.* 1983;138:375–379.
12. Handsfield HH, Jasman LL, Roberts PL, Hanson VW, Kothenbeutal RL, Stamm WE. Criteria for selective screening for *Chlamydia trachomatis* in women attending family planning clinics. *JAMA.* 1986; 255:1730–1734.
13. Johnson BA, Poses RM, Fortner CA, Meier FA, Dalton HP. Derivation and validation of a clinical diagnostic model for chlamydial cervical infection in university women. *JAMA.* 1990;264:3161–3170.
14. Rothenberg R. The geography of syphilis: a demonstration of epidemiologic diversity. In: Morriset R, Kurstak E, eds. *Advances in Sexually Transmitted Diseases.* Utrecht, Netherlands: VNU Science Press; 1986: 125–133.
15. Schmid GP, Sanders LL, Blount JH, et al. Chancroid in the United States: reestablishment of an old disease. *JAMA.* 1987; 258:3265–3265.
16. Koutsky LA, Galloway DA, Holmes KK. Epidemiology of genital human papillomavirus infection. *Epidemiol Rev.* 1988;10: 122–163.
17. Corey L. Genital herpes. In: Holmes KK, Mardh PA, Sparling PF, et al, eds. *Sexual Transmitted Diseases.* 2nd ed. New York, NY: McGraw-Hill; 1990:391–414.
18. Syrjanen K, Vayrynen M, Castren O, et al. Sexual behavior of women with human papillomavirus (HPV) lesions of the uterine cervix. *Br J Vener Dis.* 1984;60:243–248.
19. Greenberg JB, Aral SO, Magder L. Age at sexual debut: A marker for risky sexual behavior. Poster presentation at the Society of Behavioral Medicine 12th Annual Conference, March 1991, Washington, DC.
20. Dan B. Sex and the singles' whirl: the quantum dynamics of hepatitis B. *JAMA.* 1986; 256:1344.
21. Gagnon JH, Lindenbaum S, Martin JL, et al. Sexual behavior and AIDS. In: Turner CF, Miller HG, Moses LE, eds. *AIDS Sexual Behavior and Intravenous Drug Use.* Washington, DC: National Academy Press; 1989:73–185.
22. Centers for Disease Control. Premarital sexual experience among adolescent women, United States, 1970–1988. *MMWR.* 1991;39:929–932.
23. Aral SO. Sexual behavior and risk for sexually transmitted infections. *STD Bull.* 1991;10(5):3–10.
24. Forrest JD, Singh S. The sexual and reproductive behavior of American women, 1982–1988. *Fam Plann Perspect.* 1990; 22(5):206–214.
25. Seltzer VL, Rabin J, Benjamin F. Teenagers' awareness of the acquired immunodeficiency syndrome and the impact on their sexual behavior. *Obstet Gynecol.* 1989; 74(1):55–59.
26. Smith TW. Adult sexual behavior in 1989: Number of partners, frequency of intercourse, and risk of AIDS. *Fam Plann Perspect.* 1991;23(3):102–107.
27. Judkins DR, Mosher WD, Botman S. National Survey of Family Growth: design, estimation, and inference. *Vital Health Stat [2].* 1991; no. 109.
28. Aral SO, Cates W. The multiple dimensions of sexual behavior as risk factor for sexually transmitted disease: the sexually experienced are not necessarily sexually active. *Sex Transm Dis.* 1989;12:173–177.
29. Jessor R, Costa F, Jessor SL, Donovan JE. Time of first intercourse: a prospective study. *J Pers Soc Psychol.* 1983;44:608–626.
30. Mosher WD, McNally JW. Contraceptive use at first intercourse: United States, 1965–1988. *Fam Plann Perspect.* 1991; 23(3):108–116.
31. Peters RK, Thomas D, Hagan DG, Mack TM, Henderson BE. Risk factors for invasive cervical cancer among Latinas and non-Latinas in Los Angeles County. *JNCI.* 1986;77:1063–1077.
32. Harris RW, Brinton LA, Cowdell RH, et al. Characteristics of women with dysplasia or carcinoma in situ of the cervix uteri. *Br J Cancer.* 1980;42:359–369.
33. UK Family Planning Research Network. Patterns of sexual behavior among sexually experienced women attending family planning clinics in England, Scotland, and Wales. *Br J Fam Plann.* 1988;14:74–82.

34. DuRant RH, Sanders JM. Sexual behavior and contraceptive risk taking among sexually active adolescent females. *J Adolesc Health Care.* 1989;10:1–9.
35. Shafer M-A, Boyer CB. Psychosocial and behavioral factors associated with risk of sexually transmitted diseases, including human immunodeficiency virus infection, among urban high school students. *J Pediatr.* 1991;119:826–833.
36. Phillips RS, Hanff PA, Wertheimer A, Aronsom MD. Gonorrhea in women seen for routine gynecologic care: criteria for testing. *Am J Med.* 1988;85:177–182.
37. Shafer MA, Beck A, Blain B, Dole P, Irwin CE Jr, Sweet R, Schachter J. *Chlamydia trachomatis:* important relationships to race, contraception, lower genital tract infection, and Papanicolaou smear. *J Pediatr.* 1984;104:141–146.
38. Herrero R, Brinton LA, Reeves WC, et al. Sexual behavior, venereal diseases, hygiene practices, and invasive cervical cancer in a high-risk population. *Cancer.* 1990; 65:380–386.
39. La Vecchia C, Franceschi S, Decarli A, et al. Sexual factors, venereal diseases, and the risk of intraepithelial and invasive cervical neoplasia. *Cancer.* 1986;58:935–941.
40. Kjaer SK, Teisen C, Haugaard BJ, et al. Risk factors for cervical cancer in Greenland and Denmark: a population-based cross-sectional study. *Int J Cancer.* 1989; 44:40–47.
41. Parazzini F, La Vecchia C, Negri E, Fasoli M, Cecchetti G. Risk factors for adenocarcinoma of the cervix: a case-control study. *Br J Cancer.* 1988;57:201–204.
42. Vener AM, Stewart CS. Adolescent sexual behavior in middle America revisited: 1970–1973. *J Marriage Fam.* 1976;38:728–735.
43. Darrow WW. Social stratification, sexual behavior, and the sexually transmitted diseases. *Sex Transm Dis.* 1979;6:228–230.
44. Morton WE, Horton HB, Baker HW. Effects of socioeconomic status on incidences of three sexually transmitted diseases. *Sex Transm Dis.* 1979;6:206–207.
45. Seidman SN, Aral SO. Race differentials in STD transmission. *Am J Public Health.* 1992; 82:1297. Letter.

The Medical Origins of Homelessness

Marilyn A. Winkleby, PhD, Beverly Rockhill, MA, Darius Jatulis, MS, and Stephen P. Fortmann, MD

ABSTRACT

In 1989 through 1990, we conducted a cross-sectional survey of 1437 homeless adults in northern California (98% response rate). Prevalences of alcohol abuse, illegal drug use, and psychiatric hospitalization when adults first became homeless were 15% to 33% lower than prevalences following homelessness. The largest differences between the homeless and a comparison group of 3122 nonhomeless adults were for psychiatric hospitalization (odds ratios [ORs] of 4.6 for men and 5.9 for women) and alcohol abuse (ORs of 2.3 for men and 4.0 for women). However, when prehomeless prevalences of addictive and psychiatric disorders were compared with prevalences among the nonhomeless, absolute differences were no greater than 12%. (*Am J Public Health.* 1992;82:1395–1398)

Introduction

A consistent sociodemographic profile of homeless adults has emerged over the last decade indicating that the homeless are predominantly men in their 30s who have completed high school,[1,2] have few or no dependents,[3,4] and are long-term residents of the cities in which they reside.[1,2] Research on substance abuse[5–8] and psychiatric disorders[8–17] among the homeless is less consistent and remains the subject of considerable debate[8,18,19] because of two conceptual and analytic problems. First, with some exceptions,[15] addictive and psychiatric disorders that precede homelessness have not been distinguished from those that are a consequence of homelessness, thereby limiting the opportunity to explore causal hypotheses and to evaluate the accretion of disorders following homelessness. Second, many studies have not used comparative data on nonhomeless populations, thereby hindering the evaluation of the extent to which the homeless exhibit an excessive prevalence of medical disorders.

Methods

We conducted a cross-sectional survey of adults from the three National Guard armories in Santa Clara County, California, that provide approximately half of all shelter beds for adults in the county.[20,21] The largest armory (with a capacity of 300 persons) is located in downtown San Jose, the 13th largest city in the United States. The armories admit any adult without children, allow registration throughout the night, and impose no limits on the length of stay (the average stay per guest is 15 nights). Recruitment took place every night from November 26, 1989, through March 31, 1990. All persons were invited to participate on the first night they registered and were administered a 58-item questionnaire by nurses and health interviewers. To ensure that no one was interviewed twice, a master file was kept of names, ages, and interview status.

Demographic questions were adapted from the 1980 US census. Items on substance abuse were selected from the National Institute of Mental Health (NIMH) Diagnostic Interview Schedule (DIS), a structured questionnaire designed for lay interviewer administration.[22] The DIS questions examining excessive alcohol intake addressed both perceptual ("Did you ever think you were an excessive drinker?) and quantitative ("Did you ever consume

The authors are with the Stanford Center for Research in Disease Prevention, Stanford University School of Medicine, Palo Alto, Calif.

Requests for reprints should be sent to Marilyn A. Winkleby, PhD, Stanford Center for Research in Disease Prevention, Stanford University School of Medicine, 1000 Welch Rd, Palo Alto, CA 94304-1885.

This paper was submitted to the Journal June 10, 1991, and accepted with revisions June 8, 1992.

seven or more beers, drinks, or glasses of wine every day for two weeks or more?") aspects of drinking. The DIS question on drug use asked whether respondents had ever used illegal drugs (including opiates, cocaine, and marijuana) every day for 2 weeks or more. The psychiatric hospitalization question, selected from the NIMH Epidemiologic Catchment Area Study questionnaire,[23] asked whether respondents had ever been admitted overnight to a hospital or treatment program for family or personal problems, emotional problems, or problems with drugs or alcohol. These questions were adapted to distinguish whether addictive and psychiatric disorders preceded or followed the respondent's first episode of homelessness.

We compared our data with data for adults from three nonhomeless populations: (1) the Epidemiologic Catchment Area Study (the California DIS sample, n = 3122)[24]; (2) the 1980 California census, representing the general domiciled population[25]; and (3) a subset from the 1980 California census, representing a local impoverished domiciled population.[26] The last group is a census bureau statistical subarea and includes 64 430 adults from 23 low-income San Jose census tracts. Analyses comparing the homeless with the California DIS sample were stratified by sex and directly standardized for marital status because of large differences in prevalence by sex and disproportionate numbers of unmarried respondents among the homeless.[27]

To assess the reliability of respondents' answers, 35 individuals were randomly selected and reinterviewed at the conclusion of the study. All variables showed a high degree of agreement in the test/retest analysis; 100% for age first homeless and illegal drug use, 91% to 97% for adverse childhood events, and 86% to 89% for alcohol use and psychiatric hospitalization.

Results

Response Rate

Of the 1464 adults who registered at the armories, 1268 men and 169 women (n = 1437) completed the survey. The response rate was 98%, with no significant differences between respondents and nonrespondents by age, sex, or race ($P > .05$).

Homeless Characteristics

Most respondents first became homeless during young adulthood (47% before age 30, 78% before age 40). Forty-four percent had been homeless more than once, and 39% had spent more than 1 year without shelter (cumulative lifetime episodes). During the month before registering at the armories, 63% were literally homeless (i.e., their usual sleeping place was outside or in deserted buildings, in vehicles, or in shelters).

TABLE 1—Sociodemographic Profile of the Homeless Sample Compared with Two Nonhomeless Samples

	Homeless, %			Nonhomeless, %	
	Men (n = 1268)	Women (n = 169)	Total (n = 1437)	Impoverished Sample	California Sample
Current age, y					
15–24	14	22	15	32	24
25–34	36	37	36	23	23
35–44	31	24	30	11	15
45–54	11	8	11	10	13
55–64	6	6	6	10	12
≥65	2	2	2	15	13
Race					
Caucasian	47	51	48	47	67
Hispanic	27	15	25	41	19
African American	21	21	21	3	8
Asian	1	2	1	8	6
Other	4	10	5	1	1
Education[a]					
<High school	32	33	32	47	25
High school	38	38	38	23	27
Some college	23	20	23	17	23
≥4 years college	7	9	7	13	24
Placed in foster care before age 18	10	17	11	...	...
Sexually abused before age 18	6	33	9	...	...
Physically abused before age 18	13	28	14	...	...

[a]Includes those 25 years of age and older. When limited to those born in the United States, the percentages for the homeless population are 26% <high school, 41% high school, 25% some college, and 7% ≥4 years of college.

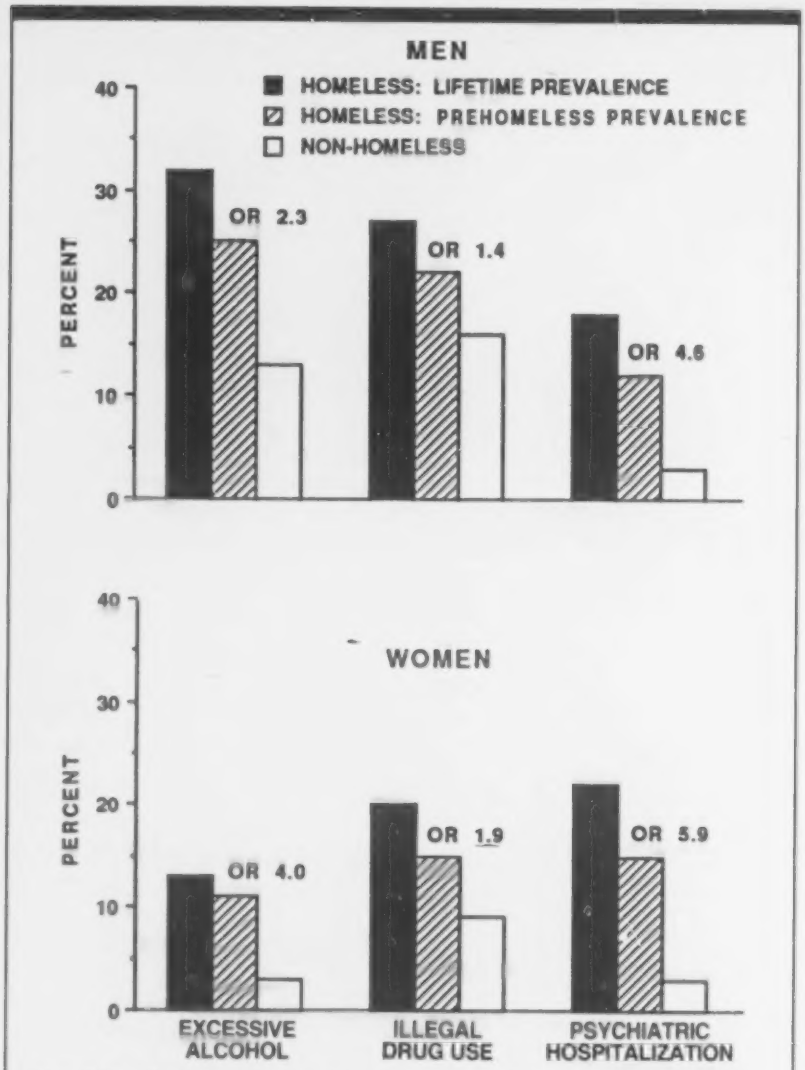
Sociodemographic Profile

Two thirds of the homeless were between the ages of 25 and 44 years (mean age = 36 years); virtually none were 65 or older (Table 1). Compared with the two nonhomeless census samples, African Americans were overrepresented and Asians were underrepresented among the homeless. The higher proportion of African Americans is consistent with their overrepresentation among the domiciled poor (e.g., 30% of African Americans live below the poverty level, compared with 9% of Caucasians),[25] and suggests that they may be more vulnerable to homelessness when affordable housing is scarce.[15] The small percentage of homeless Asians, however, is unexpected, given the recent influx of very poor Southeast Asian refugees to the area.

Sixty-eight percent of the homeless had completed at least high school and 30% had attended college. These levels of schooling increased when the analysis was restricted to respondents born in the United States (of whom 73% had completed high school and 32% had attended college). Compared with the nonhomeless impoverished sample, the homeless were less likely to have dropped out of high school; however, compared with Californians overall, the homeless had less education, especially at the college level.

Adverse Childhood Events

Ten percent of the men and 17% of the women had been removed from their parents and placed in foster care before the age of 18 (46% of these respondents were removed before age 10 and 54% at age 10 or older). Two general questions ("Were you ever physically abused before age 18?" and "Were you ever sexually



Note. Odds ratios (OR) compare prehomeless prevalences to nonhomeless prevalences.

FIGURE 1—Prevalences of excessive alcohol intake, illegal drug use, and psychiatric hospitalization among the homeless and the nonhomeless (California Diagnostic Interview Schedule sample).

abused before age 18?") indicated that 6% of the men and 33% of the women had been sexually abused and that 13% of the men and 28% of the women had been physically abused. Thirty-eight percent of those reporting physical abuse and 39% of those reporting sexual abuse had also been placed in foster care.

Adult Addictive and Psychiatric Disorders

Prevalences of excessive alcohol intake (determined on the basis of the quantitative alcohol question), illegal drug use, and psychiatric hospitalization before homelessness were 15% to 33% lower than lifetime prevalences (Figure 1). When comparing addictive and psychiatric disorders of the homeless before their initial loss of shelter (prehomeless prevalences) with those of the nonhomeless, the largest differences in prevalences were for alcohol intake (odds ratios [ORs] of 2.3 for men and 4.0 for women) and psychiatric hospitalization (ORs of 4.6 for men and 5.9 for women). ORs were 1.7 for both men and women when the perceptual alcohol question was used. Illegal drug use showed the smallest difference between the homeless before their initial loss of shelter and the nonhomeless (ORs = 1.4 for men and 1.9 for women). Despite the higher prevalences of addictive disorders and psychiatric hospitalization among the homeless than among the nonhomeless, absolute differences were no larger than 12 percentage points.

Forty-eight percent of the men and 63% of the women entered homelessness with no adult medical disorders (alcohol or drug abuse or psychiatric hospitalization); and 80% of the men and 59% of the women entered homelessness with no history of adverse childhood events (placement in foster care or sexual or physical abuse). To test which adverse childhood events and adult medical disorders were most strongly related to total time homeless, these factors were entered simultaneously into an ordinal logistic regression model[28] that adjusted for years of education. Excessive alcohol use as an adult and physical abuse during childhood were the factors most strongly related to length of time homeless (*P*s = .003 and .006, respectively).

Discussion

Although the prevalences of addictive and psychiatric disorders were higher among the homeless than among the nonhomeless, the differences may not be as great as previously reported. The prevalences of substance abuse and psychiatric hospitalization before homelessness (the prevalence rates most appropriate for studying causal hypotheses) were 15% to 33% lower than lifetime prevalences (the rates reported by most homeless studies). Furthermore, when the prehomeless prevalences were compared with prevalences among the nonhomeless, absolute differences were no greater than 12%.

Limitations

The following limitations should be considered. First, our findings are based on a cross-sectional study design rather than on a more definitive prospective design. Second, our assessments of substance abuse and psychiatric hospitalization are based on several standardized questions, possibly resulting in an underestimation of adult medical disorders in both the homeless and nonhomeless comparison groups. Investigators have, however, validated 3 of the 23 DIS questions on alcohol abuse against the Diagnostic and Statistical Manual of Mental Disorders, third edition (DSM-III), in a California homeless population[29] and found high sensitivity (94%), specificity (85%), and predictive value (91%). Third, although we used psychiatric hospitaliza-

tion as a measure of mental disorders because of its frequent use in past studies[2,30] and because it allowed for a valid comparison among the homeless and nonhomeless samples, we support the views of previous investigators who have emphasized the importance of structured clinical interviews to diagnose mental disorders,[10,18,22] and we recognize that psychiatric hospitalization does not serve as a direct marker for mental illness. In addition, the NIMH question on psychiatric hospitalization does not distinguish hospitalization for psychiatric problems from hospitalization for substance abuse problems. Finally, the generalizability of our results is unknown, because we sampled only shelter users. However, data we obtained from a telephone survey of homeless shelter and soup kitchen managers in the county (n = 15) and from a pretest survey of homeless men living in a riverbank encampment (n = 20) suggest that there are not important differences in prevalences of addictive and psychiatric disorders between shelter users and non–shelter users. Nevertheless, our sample, which included only adults without children, may be unrepresentative of homeless populations from other geographic areas, especially those from urban inner cities.

Consistency of Findings

Our sample appears quite similar to other homeless samples in regard to age (mean age of our respondents was 36 years, the same age as averaged over 23 studies)[2] and ethnic representation (49% of our respondents were minorities, compared with 44% of the homeless populations from the National Shelter Survey).[31] The lifetime prevalences of medical disorders we found are also in agreement with those found in past work, with approximately one in three persons reporting excessive alcohol intake and one in four to five reporting illegal drug use or psychiatric hospitalization.[2,30] In addition, the prevalences of medical disorders preceding homelessness found here mirror those reported by a previous study that sampled first-time shelter users and included a subsample of men interviewed on the first night they were homeless (n = 56).[15] In that study, the prevalence of psychotic disorders was lowest in the newly homeless and increased with length of time homeless.

Conclusions

Our results suggest that the homeless are less deviant in terms of social, addictive, and psychiatric pathology than has previously been proposed. Furthermore, our findings illustrate the need to examine disorders antecedent to the loss of shelter and to present comparative data on nonhomeless populations. □

Acknowledgments

This research was supported by grants from The Valley Foundation, the Community Foundation of Santa Clara County, the John Arrillaga Foundation, and the Komes Foundation.

This paper was presented at the 119th annual meeting of the American Public Health Association, Atlanta, Ga, November 10–14, 1991.

We are indebted to Barry Del Buono of the Emergency Housing Consortium and John Hall of the Health Care for the Homeless Project for access to the study population; to Byron Brown, Jr, for statistical consultation; and to Michael Fischetti, Andrew Zolopa, Erica Frank, and Sanford Dornbusch for their advice on the manuscript.

References

1. Committee on Health Care for Homeless People, Institute of Medicine. *Homelessness, Health, and Human Needs.* Washington, DC: National Academy Press; 1988.
2. Rossi PH. Down and out in America: the origins of homelessness. Chicago, Ill: University of Chicago Press; 1989.
3. Piliavin I, Sosin M, Westerfelt H. Conditions contributing to long-term homelessness: an exploratory study. Madison, Wis: Institute for Research on Poverty; 1987. Discussion Paper No. 853-87.
4. Sosin MR, Colson P, Grossman S. *Homelessness in Chicago: Poverty and Pathology, Social Institutions and Social Change.* Chicago, Ill: University of Chicago Press; 1988.
5. Wright JD, Weber E. *Homelessness and Health.* New York, NY: McGraw-Hill; 1987.
6. Koegel P, Burnam A. Alcoholism among homeless adults in the inner city of Los Angeles. *Arch Gen Psychiatry.* 1988;45: 1011–1018.
7. Lubran B. Alcohol and other drug problems among the homeless population. *The Nation's Health.* September 1989;23–25.
8. Fischer PJ, Breakey WR. The epidemiology of alcohol, drug, and mental disorders among homeless persons. *Am Psychol.* 1991;46:1115–1128.
9. Bassuk EL. The homelessness problem. *Sci Am.* 1984;251:40–45.
10. Robertson MJ. Mental disorder among homeless persons in the United States: an overview of recent empirical literature. *Admin Ment Health.* 1986;14:14–27.
11. Roth D, Bean GJ. New perspectives on homelessness: findings from a statewide epidemiological study. *Hosp Community Psychiatry.* 1986;37:712–719.
12. Snow DA, Baker SG, Anderson L, Martin M. The myth of pervasive mental illness among the homeless. *Soc Probl.* 1986;33: 406–423.
13. Lamb HR. The homeless mentally ill. In: *Homelessness: Critical Issues for Policy and Practice.* Boston, Mass: The Boston Foundation; 1987:33–37.
14. Breakey WR, Fischer PJ, Kramer M, et al. Health and mental health problems of homeless men and women in Baltimore. *JAMA.* 1989;262:1352–1357.
15. Susser E, Struening EL, Conover S. Psychiatric problems in homeless men. *Arch Gen Psychiatry.* 1989;46:845–850.
16. Koegel P, Burnam MA. Problems in the assessment of mental illness among the homeless: an empirical approach. In: Robertson MJ, Greenblatt M, eds. *Homelessness: The National Perspective.* New York, NY: Plenum; 1990.
17. Ritchey FJ, La Gory M, Fitzpatrick KM, Mullis J. A comparison of homeless, community-wide, and selected distressed samples on the CES-depression scale. *Am J Public Health.* 1990;80;1384–1386.
18. Susser E, Conover S, Struening EL. Problems of epidemiologic method in assessing the type and extent of mental illness among homeless adults. *Hosp Community Psychiatry.* 1989;40:261–265.
19. Winkleby MA. Comparison of risk factors for ill health in a sample of homeless and nonhomeless poor. *Public Health Rep.* 1990;105:404–410.
20. Winkleby MA, White R. Homeless adults without apparent medical and psychiatric impairment: onset of morbidity over time. *Hosp Community Psychiatry.* 1992; in press.
21. Winkleby MA, Fleshin D. Are homeless veterans at higher risk for physical, addictive, or psychiatric disorders than homeless non-veterans? *Public Health Rep.* 1992; in press.
22. Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule. *Arch Gen Psychiatry.* 1981;38:381–389.
23. Regier DA, Myers JK, Kramer M, et al. The NIMH Epidemiologic Catchment Area program: historical context, major objectives, and study population characteristics. *Arch Gen Psychiatry.* 1984;41:934–941.
24. Burnam MA, Hough RL, Escobar JI, et al. Six-month prevalence of specific psychiatric disorders among Mexican Americans and non-Hispanic whites in Los Angeles. *Arch Gen Psychiatry.* 1987;44:687–694.
25. *General Social and Economic Characteristics. Part 6, California, Section 1.* Washington, DC: US Bureau of the Census; 1982.
26. *San Jose Neighborhood Statistics Program 1980 Census Data.* San Jose, Calif: Department of City Planning; 1983.
27. Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York, NY: John Wiley & Sons Inc; 1981.
28. McCullagh P, Nelder JA. Generalized linear models. 2nd ed. New York, NY: Chapman & Hall; 1989.
29. Vernez G, Burnam A, McGlynn EA, Trude S, Mittman BS. Review of Califor-

nia's Program for the Homeless Mentally Disabled. Santa Monica, Calif: RAND Corporation; 1988.
30. Caton CLM. Homelessness in historical perspective. In: Caton CLM, ed. *Homeless in America.* New York, NY: Oxford University Press; 1990:1–18.
31. *A Report to the Secretary on the Homeless and Emergency Shelters.* Washington, DC: US Dept of Housing and Urban Development, Office of Policy Development and Research; 1984.

The Robert Wood Johnson Foundation's Call for Proposals on Health Care Financing

The Robert Wood Johnson Foundation is requesting proposals for research, demonstration, and evaluation projects that will produce information useful for policymakers regarding major changes in health care financing. This call for proposals is intended to stimulate projects that: (1) examine the effects of current mechaniams for, and proposed major changes in, the financing of health services on health care costs, access, and quality; and (2) develop and test new ways to finance care that have the potential to improve access to more affordable health services. Up to $12 million will be made available over 3 years for this initiative. Ultimately, the foundation hopes the results of these projects will be timely and useful for policymakers, public and private.

To be considered favorably, proposals must have the clear objective of examining major issues in health care financing and their effects on costs, access, or quality. Proposals will be assessed using the following criteria:

- the policy significance of the health care financing mechanism being evaluated or tested;
- the timeliness of the project for informing policy or practice;
- the quality and availability of data to be used, and the strength of the proposed methodology (depending on the focus of the project, primary data collection may be supported or the use of secondary data may be acceptable);
- the uniqueness of the project;
- the applicant's experience and qualifications for conducting the proposed project and the time commitment of key project staff who have the skills and experience to perform the various operations and analytic tasks required.

Preference will be given to applicants that are public agencies or are tax-exempt under Section 501(c)(3) of the Internal Revenue Code and are not private foundations as defined under Section 509(a).

Letters of intent will be accepted under this solicitation through May 1995 or until authorized funds are awarded, whichever comes earlier. For more information, contact Alpha Center, attn: HCFO, Suite 1100, 1350 Connecticut Ave NW, Washington, DC 20036; tel. (202) 296-1818; fax (202) 296-1825.